### Full Title:
Inferring The Optimal Number Of Dockless Shared Bike In A New Area By Applying The Gradient Boosting Decision Tree Model

### Abstract:
This paper applies the gradient boosting decision tree (GBDT) model which is a widely used machine learning algorithm, to provide a case study to infer the overall optimal number for the dockless shared bike (DSB). Variables from the built environment, socio-demographic, transportation infrastructures, and public transport networks are considered in the paper. By analyzing DSBs' spatial distribution and relevant features from two similar cities (Nanjing, Hangzhou) in China which have experienced same time phrases: First, the DSB experienced a fast growth within the city’s core areas; Second, the private operation companies put a number of DSB in the market which went beyond the transportation infrastructure capacity. Relevant policies published after that in order to control the total quantity on the street. Third, after the same period of adjustment where both the users’ demand and the city’s ability to supply have been balanced. The case study results indicate that, ideally, around 24.3 thousand shared bikes can be implemented within the Suzhou’s central district. The model in this paper provides a general transfer learning methodology in inferring the optimal number of DSB that can be adopted to the new city area.

### Manuscript Classifications:
Operations and Traffic Management; Bicycle Transportation ANF20; Bikesharing

### Manuscript Number:

### Article Type:
Presentation and Publication

### Order of Authors:
- Dong Xiao
- Tianqi Gu
- Inhi Kim
- Yuanqiu Bao
Inferring The Optimal Number Of Dockless Shared Bike In A New Area By Applying The Gradient Boosting Decision Tree Model

Dong Xiao
Institute of Transport Studies, Civil Engineering Department
Monash University, VIC, Australia 3800
Email: dong.xiao@monash.edu

Tianqi Gu
Institute of Transport Studies, Civil Engineering Department
Monash University, VIC, Australia 3800
Email: Tianqi.gu@monash.edu

Inhi Kim, Ph.D., Corresponding Author
Institute of Transport Studies, Civil Engineering Department
Monash University, VIC, Australia 3800
Email: inhi.kim@monash.edu

Yuanqiu Bao
Director
Planning Department
CCDI Group, Suzhou, China 215123
Email: bao.yuanqiu@ccdi.com.cn

Word Count: 5,162 words + 2 tables*250 = 5,662 words

Submitted [August 1]
ABSTRACT
This paper applies the gradient boosting decision tree (GBDT) model which is a widely used machine learning algorithm, to provide a case study to infer the overall optimal number for the dockless shared bike (DSB). Variables from the built environment, socio-demographic, transportation infrastructures, and public transport networks are considered in the paper. By analyzing DSBs’ spatial distribution and relevant features from two similar cities (Nanjing, Hangzhou) in China which have experienced same time phrases: First, the DSB experienced a fast growth within the city’s core areas; Second, the private operation companies put a number of DSB in the market which went beyond the transportation infrastructure capacity. Relevant policies published after that in order to control the total quantity on the street. Third, after the same period of adjustment where both the users’ demand and the city’s ability to supply have been balanced. The case study results indicate that, ideally, around 24.3 thousand shared bikes can be implemented within the Suzhou’s central district. The model in this paper provides a general transfer learning methodology in inferring the optimal number of DSB that can be adopted to the new city area.

Keywords: Dock-less shared bike, transfer learning, feature extraction, gradient boosting decision tree
INTRODUCTION
The bike-sharing program is firstly emerged in Europe over 50 years ago which defined as the network of public use bicycles distributed around a city that allows people to rent or borrow bikes for a relatively short duration (7). As a cost-saving transfer option to provide a solution to the ‘first mile’ and ‘last mile’ problems, the shared bikes can fill the gaps between the existing transportation networks. This program not only able to enhance the utilization of other public transit modes, moreover, sufficiently reduce the ridership, which helps release the traffic congestion burden within the city.

With the rapid development of the shared mobility service, start from 2016, the dockless bike-sharing programs (DBSPs) emerged, also known as the free-floating bicycle sharing system, experienced a dramatically fast growth led by private entities within the major cities in China. Most of the dockless shared bikes (DSB) nowadays are believed as the ‘fourth generation.’

Compared with the ‘third generation’ that uses smart technology (mobile-phone, mag-stripe cards, smart-cards or codes) to unlock the bicycle from the stations (2), it is a new form of development of public bikes offering self-positioning system, intelligent terminal scanning and unlocking technology but completely eliminates stations and docks (3). Also, it brings tremendous vitality with the advantages of flexibility and high density. As of 2018, the volume of DSB ridership has reached more than 70 million at the top peak hour time point with 23 million DSB number involving in 200 cities in China (4), indicating the largest DSB market in the world and the third-largest public transportation mode except for bus and subway in China (5).

Ofo and Mobike, as the two start-up companies, initiates an innovative generation of fully dockless bike-sharing services in China (6). The unexpected venture capital market competition between the DSB operation companies with the unclear user demand estimation and unbalanced resource allocation has caused various problems. Due to a substantial number of DSB has been put into the market, and those bikes can be distributed randomly since lacking parking piles and geo-fence restrictions. Thus, public space sometimes can be occupied, and this issue especially severe in cities with increasingly limited land resources (7). In the central areas of the city, many shared bikes parking on the sidewalk block pedestrians and non-motor vehicles (4). The disordered users’ parking behaviors has brought chaos to the utilization of the public open space and sometimes may affect other transportation modes negatively.

In order to improve the city management and to help avoid unnecessary expenditure for the municipal governments, the bike-sharing program operation entities and other stakeholders, not only the demand should be estimated, but also the ideal number of DSB that the environment is capable of bearing is essential to the design of the system. Especially it can be vital to the places where have not been implemented yet. The planning authority in Beijing roughly calculated the infrastructure capacity previously by measuring the available sidewalk space and giving the maximum occupancy to the DSB, which remains insufficient. However, inferring the optimal number of DSB to a new city area is more challenging as the following problems cause it: (i) no direct historical data available in the area which no previous prediction models are feasible to cope with the users’ demand. (ii) the spatial correlation of the DSB distribution remains unrevealed as it can be dynamic in a short or a long-time period.

Most existing research on bike-sharing system focus on the station-based bike-sharing system, which is restrained by the location of the dock station (5) at the earlier stage. Tradition models are applied for the travel demand prediction (8; 9), such as using Latent Demand Score Method or disaggregated model to estimate. Some studies try to use either the bike-sharing travel
pattern or the trip frequency to reveal the influential factors that can affect the user’s demand (10). The network modeling like rebalancing problem is another primary subject that often considered due to the limited docking racks available at different stations (11; 12).

As the DBSP is a trending travel mode with different user preferences, still very few numbers of studies with different methods have been carried out to analyze the travel demand of the DBSP. The learning-based model: long short-term memory (LSTM) is firstly applied in DSBP research conducted in Nanjing, China (10). More models are introduced with emerging technologies in deep learning technology. The recurrent neural network (RNN) and its variants LSTM and gated recurrent unit (GRU) are gaining popularity among researchers (13, 14).

Considering the RNN models are unable to cope with the spatial features, the convolutional neural network (CNN) model is developed in some studies as well (15). A developed model as Convolutional LSTM is involved in the demand prediction research about DSB as well (15). The conventional artificial neural network (ANN) often practised as a demand prediction algorithm by dynamically modeling time-series data.

**Influential Factors**

Characteristics from different categories are analyzed in many papers along with the bike-sharing usage data trying to identify the influential factors that may affect users’ demand. Contrast with the dock-based shared bike system, operation companies can only adjust the shared bikes’ spatial distribution by controlling the number of dock stations and their geolocation, and the DSB can be adjusted directly. It is believed that the user behaviors can be affected by many external factors, usually categorized into two sets: the heterogeneity in travelers’ characteristic and the built environment effects. Several previous studies reveal that how the shared bikes move with the users spatially. Thus, the spatial distribution of DSB is often a reflection of their potential demand usage. For the demographic factors, many pieces of research indicated that gender and age are not found significant (16; 17). A positive relationship between education level and the DSB is revealed in the previous study (5).

The built environment is considered as the fundamental element that can profoundly affect the form of a city. Thus, the built environment should be an indispensable factor when analyzing the travel demand. The population density is found to be the most critical factors that impact the bike usage (18). Other public transit facilities like the subway station are widely analyzed from the researchers, and the results indicate that a significant positive relationship can be explored (19). The density of employment, education, recreational, and commercial is often selected in previous research (20). Others like biking infrastructures (6) and land use types (18) and weather conditions, air quality (10) are also concerned.

**Transfer Learning**

The idea of transfer learning methodology in the machine learning framework enables the transfer of knowledge from a domain with rich available data to a target domain with little data (21). This method allows us to explore the spatiotemporal knowledge of DSBP from other cities with sufficient experience from the practical operation. To the best of our knowledge, limited studies concentrated on the demand prediction for the DSB with no historical data of the target city. Only one literature is found that explicitly dealt with the transference of free-floating bikes to others (22). Factor analysis (FA) and convolutional neural network (CNN) were used in their model, where the FA is a dimensional reduction method that used to identify latent features from high-dimensional feature space, and CNN as a machine learning process is capable of capturing the spatial autocorrelation between neighboring areas.
In general, giving the fact that no suitable models can be directly adopted without the interpretation of the historical trip data or real spatial distribution of the DSB in the new area. This paper provides a macroscopic perspective view in the DSB quantity analysis and giving significant contributions to the following:

(i) Analyze the DSB distribution and multi-data source feature information from other cities to predict the overall number of DSB in the new area by concerning the city’s infrastructure capacity through the transfer learning process.

(ii) Further reveal the relationship between the DSB distribution and its potential influencing factors through the real-world data set.

(iii) Apply the Gradient Boosting Decision Tree (GBDT) model into the DSB related research. The model has been widely used in recent transportation research as it can provide high performance in prediction with markable reliability. It is firstly used in the bike-sharing system to learn the feature transformation to promote demand prediction accuracy (20). Still, the application of this model in DSB related research is minimal.

Through the case study in Suzhou central area, not only how the DSB should be distributed wisely is revealed, but also give insights for new cities that have not implemented the DSB system yet.

**METHODOLOGY**

The process of the general methodology is designed as follow:

1. Select the origin city with healthy and stable DBS systems. The cities are recommended to have similar geographical environment and development level to enhance the model’s transferability.

2. Divide the area of the reference city A and a target city C into grids with a proper size. Each of these grids is denoted as $G_{Ci} \mid i \in [1, m], i \in \mathbb{N}^+$, Where $C$ is the index of the city and $i$ is the numbering of grids in each city. The distribution of dockless shared bike in the reference city A is denoted as $D_{aj} \mid j \in [1, n], j \in \mathbb{N}^+$, and $j$ is the numbering of grids in the reference city.

3. Extract various features from all grids that store multi-sources data, in this study referred to as ‘Feature Extraction’ in the flowchart diagram. The features from four categories are selected in this case: built environment (point of interest), socio-demographic (population), transport infrastructure, and the public transit network are considered.

4. Train the GBDT model with the training data set extracted from the reference city. Here, the model input is the grid features after the standardization process, and the model output is the DBS number in this grid. The available DBSs would be used as the labeled result.

5. Predict the target city grid supply with the model trained in step 4 and new city grid features with the standard format, while the results are the predicted number of every new city grids. By adding all grids value together, the inferred optimal amount can be acquired shown as ‘Target City Inference’ in the diagram.
Xiao, Kim, Bao

![Methodology Flowchart]

**Figure 1 The Methodology Flowchart**

**Study Area Selection**

Many cities in China started limiting the dock-less bike number since its large volume has already exceeded the cities’ capacity (4). In August 2017, the Department of Transportation in Shanghai firstly published the policy to restrict any new DSB put in the market and required the private operating companies to control the number of DSB on the street. Right after Shanghai, many cities’ management authorities followed up. At least two cities in China: Nanjing and Hangzhou are explicitly selected in the paper to establish the prediction model to infer the number DSB in the case study area: Suzhou central area. Even though the operation starts dates in the two cities vary, the dates when the governments publish the policies to limit the DSB number are fallen in the same month (both in August 2017). Three areas: Suzhou, Hangzhou, and Nanjing, are all defined as the second-tier city, which shares many commons in population size and economic development.

The target city, Suzhou, is located in Jiangsu Province in China, which has experienced fast growth in population and urban expansion in the past decade. With a city over a thousand year of history, how to combine the transportation development and the protection of the city’s unique
historical features becomes a vital issue to its future sustainable development. As the Suzhou
government concerns the potential drawbacks, the dockless bike-sharing systems can bring to the
urban environment observed from other cities in China. Until now (July 2019), no private
operation companies have obtained a license to run the bike-sharing system within the city’s
limit. Suzhou has more than 46,000 docked bikes and more than 2,000 docking stations by 2018
(13). However, the drawbacks of the DSB can be avoided if the optimal number can be
accurately calculated and adequately distributed.
This methodology is founded on a proposed assumption that, after few months of adjustment
since the publication of the restriction policies, by November 2017, the number of DSB in
Nanjing and Hangzhou has been well controlled to the roughly optimal quantity number. DSB’s
spatial locations are considered ideally distributed around the area, which is influenced by the
relevant factors in the built environment, socio-demographical and the transportation facilities
and network around this area.

**Gradient Boosting Decision Tree**
Both Nanjing and Hangzhou are selected as reference cities in this case with a different
purpose. The data from Nanjing is trained in the Gradient Boosting Decision Tree (GBDT)
model for analysis, while Hangzhou’s data is treated as the test set to monitor the model’s
accuracy.

Compared with the traditional logistic regression model that can only capture linear relations,
the GBDT model can be widely used for solving all regression relation (linear/non-linear)
problems. It has advantages in dealing with variables with missing values, addressing the
multicollinearity issue with a more accurate prediction (23) and even cope with the data with a
small sample size (24). In the decision tree model, all the features are divided into small subsets
based as a structure of a tree. The target of splitting data is to make the entropy reach the zero
with all features fall into the same leaf node contain the same values.

The core of the GBDT model is that thousands of single decision tree models are contained,
and each decision tree learns the residuals of all previous trees’ conclusion. This residual is the
accumulated amount of real values after adding the predicted value $\eta_i$ (see equation 1). During
the time of the training, the lost function $L$ (see equation 2) is taken as the parameter to estimate
the residual $\eta_i$, then a new decision tree can be formed with the new residual.

$$\eta_i = \frac{\partial L}{\partial f(x)^i} y_i - f(x)^i$$

$$L = \sum (y_i - f(x)^i)^2$$

$$f(x)^i = f(x)^i + a(\nabla L)$$

where $f(x)^i$ is a set of prediction; $L$ is the error of the prediction. (25)

By using the pre-trained model from the reference city, the DSB distribution in the target city
can be inferred.

**DATA PREPARATION**
Relevant dock-less shared bike data is collected through open Mobike application-programming
interface (API), Mobike data from two cities in China are selected: Nanjing and Hangzhou. All
three cities in this study are defined as the second-tier city, which has a similar population size
and urban structure form. As the central city area is more likely to be affected by the
outnumbered DSB, only the Mobike data in the central area of the two reference cities are
selected. The dataset contains Mobike DSB location information in the midnight timepoint of
11/13/2017. This date is a typical working day (Wednesday), and no special events or severe
weather happened before or after this time. The DSBs are at the most logical point at the midnight time after experienced the peaks during the day time and relevant spatial relocation operated by stakeholders. Outliers and duplicated data are excluded resulting that more than 130,000 DSBs are abstracted from Hangzhou and Nanjing.

Figure 2 Dock-less shared bike distribution heat maps (Left: Nanjing; Right: Hangzhou)

The subject of the model is defined as 1km*1km grids, which helps analysis the relevant features within the study areas. Around 1,748 and 1,188 grids are created around Nanjing and Hangzhou, respectively. About the target city Suzhou, the same process is conducted among the area, and about 2,394 grids are created. After contrast with the real-world satellite images, the grids fully contain unique terrains (e.g., mountain, river, and lakes) are removed. After the clearance of the data, the number of grids that can be input into the model is explicit. The cleaned grids are 948 in Nanjing and 816 in Hangzhou while the subject in Suzhou is 2,005 grids.

Figure 3 Grids subject created in the reference cities and the target city
Built Environment. As the previous literature has indicated that the number of POIs reflects its prosperity, thus can profoundly affect the amount of DSB distributed within the area. The features fall in the built environment set are extracted through Gaode Map API. Only POI quantity is calculated in each grid in both three cities. As commuting takes a large percentage of the total ridership \((4)\), the residential area and workplaces, education places (only level upper than elementary school is considered as no children under 12 are allowed to use the shared bikes) are firstly crawled. Other sites, like shopping and life service places, are also regarded as variables that affect the DSB spatial distribution, are input in the model.

Socio-demographic. A place with more residents attends to have more potential demand usage, which can profoundly influence the spatial distribution of DSB \((18)\) as other factors like gender and age. Since all the grids have the same size, only the total population number in each grid is abstract, which assumes the density within the area.

Road Network. The DSB’s spatial movement cannot be isolated with the transportation network. The roadway design, especially the length of the road network within each grid, is included. The area with a higher density of the road network means more capacity in bearing DSB, which can lead to more DSB distribution. Also, along with the major public transport networks’ length: Subway and bus are set as attributes in the model.

Transport Facility. The DSB has a function in connecting users to the other transportation modes. Thus, the bus stops and subways stations are playing vital roles in this park and ride mode, which attracts several DSB distributed around with the short distance. The parking space is also taken into consideration as an essential influencing factor in DSB distribution.

### TABLE 1 Detailed features in data sets

<table>
<thead>
<tr>
<th>Feature Set</th>
<th>Features</th>
<th>City</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Nanjing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Built Environment (POIs)</td>
<td>Residential Area</td>
<td>19.33</td>
</tr>
<tr>
<td></td>
<td>Work Places</td>
<td>60.49</td>
</tr>
<tr>
<td></td>
<td>Life Service</td>
<td>43.16</td>
</tr>
<tr>
<td></td>
<td>Education</td>
<td>14.49</td>
</tr>
<tr>
<td></td>
<td>Restaurant</td>
<td>48.59</td>
</tr>
<tr>
<td></td>
<td>Tourism Spot</td>
<td>4.29</td>
</tr>
<tr>
<td></td>
<td>Hotel</td>
<td>5.71</td>
</tr>
<tr>
<td></td>
<td>Shopping</td>
<td>67.02</td>
</tr>
<tr>
<td>Socio-demographic</td>
<td>Population</td>
<td>4903.85</td>
</tr>
<tr>
<td>Road Network</td>
<td>Road Network</td>
<td>7471.41</td>
</tr>
<tr>
<td></td>
<td>Bus Route</td>
<td>6159.62</td>
</tr>
<tr>
<td></td>
<td>Subway Route</td>
<td>249.41</td>
</tr>
<tr>
<td>Transport facility</td>
<td>Bus Stop</td>
<td>2.65</td>
</tr>
<tr>
<td></td>
<td>Subway Station</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>Parking Spot</td>
<td>18.50</td>
</tr>
</tbody>
</table>

Note: Std. = Standard deviation.
After the clearance of data, using the ArcGIS (Geographical Information System) software to correlate multi-source feature data, number of DSB and the grids, the spatial geographic feature information (quantity or length) is calculated in each grid. Then the formatted data can be input into the GBDT model for analysis. The only missing data part here is the ‘number of DSB’ in subjects of Suzhou, which are the model’s output.

RESULTS

Model Training
While Nanjing is chosen as the training set to build the model, and Hangzhou is treated as the test set to verify the accuracy of the model.

Through the ongoing learning adjustment of the GBDT model, the data set in Nanjing provides an outcome with a trained model. The standardized data from Hangzhou is input into this pre-trained model. The result indicates that a total number of 60,742 DSB is predicted within the area by the given information. Compared with the current number of DSB taking up to 52,006, the model provides the expected value with a 16.8% MAPE.

Generalization
Since the data collection methods can be different in regions, the type of information available varies. A robustness test of the model is proposed, as to see the model prediction performance with incomplete data.

The output of the GBDT model can provide the index of the feature importance. It is widely accepted that the order of importance means how important the feature is in terms of determining the target variable (number of DSB) (26). The ranking results are illustrated in Figure 4 after the training of the Nanjing’s data.

![The feature importance ranking](image)

The feature importance ranking indicates that the areas with higher road network density and more workplace density tend to have the most significant influence on the DSB distribution. While the subway stations location and the hotel density have the lowest impact among all the features.
in this case. The GBDT model, with some particular features removed from the model input, is retrained. The results are shown in Table 2 below:

**TABLE 2 Robustness test**

<table>
<thead>
<tr>
<th>Importance Ranking</th>
<th>Feature</th>
<th>Prediction (Hangzhou)</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>Road Network Length</td>
<td>60,488</td>
<td>16.3%</td>
</tr>
<tr>
<td>#2</td>
<td>Workplace</td>
<td>60,716</td>
<td>16.7%</td>
</tr>
<tr>
<td>#3</td>
<td>Bus Length</td>
<td>60,058</td>
<td>15.5%</td>
</tr>
<tr>
<td>#4</td>
<td>Population</td>
<td>60,428</td>
<td>16.2%</td>
</tr>
<tr>
<td>#5</td>
<td>Parking</td>
<td>60,176</td>
<td>15.7%</td>
</tr>
<tr>
<td>#6</td>
<td>Life Service</td>
<td>60,379</td>
<td>16.1%</td>
</tr>
</tbody>
</table>

By removing the features ranked at the top six from the model, respectively, the new pre-trained models are implemented to predict the overall DSB number in Hangzhou. The MAPEs are changing within 1.3% of differential from the model with all feature included. This unstable fluctuation falls within an acceptable range. The result illustrates that the model has good robustness, which has the stability to be applied to other areas with available differentiated data.

**Inferred DSB Number**

After the training process to build the model, the grids with unique variable attributes in the Suzhou area are input into the model with the missing data in DSB distribution. The outcome of the model assigns every single grid with the predicted average number of DSB, which is mapped as showing in Figure 5:

**Figure 5 The spatial distribution of DSB number in Suzhou central area**
Each grid contains a unique value of DSB. By adding all the predicted values from each grid, the total predicted number of DSB in the Suzhou central area is 145,501. The MAPE remains the same with Hangzhou as 16.8%. The model input only considers the Mobike data at the very beginning, while other brands of DSBP are excluded. By giving the fact that the Mobike company takes about 60% of the total private-owned DSB market in most the cities, the overall optimal number of DSB that can be put into the whole area is supposed to be around 242,500 in total to match with the ideal supply with other cities.

CONCLUSION AND DISCUSSION

The DSBP is booming in China and has become one of the essential travel modes. While the majority of the operating entities are private-owned, and they were purely seeking maximum profit by nature. Thus, the quantity of the dockless shared bike has experienced disorderly growth for market competition in many cities at the early stage. By all kinds of regulations and policies to control, the number has been sized in a relatively optimal amount not exceed the transport infrastructure capacity. A challenging issue for other cities which do not have the DBSP yet but plan to deploy it in the future is how to avoid this unnecessary phase by setting a limitation at the beginning.

The study investigated the DSB’s spatial distribution and their potential, influential factors within the city of Nanjing and the city of Hangzhou and developed a forecasting model by using the gradient boosting decision tree. By applying this model in the central Suzhou as the case study area, an optimal number of DSB within this area are inferred: The total number of the dockless shared bike can be introduced around 242,500 with notable 16.8% MAPE.

The application of the model has shown its generalization ability which provides a solution to the estimation of the overall number which may be further applied in a general case to places like Suzhou without any real-world distribution, or usage data can be collected and analyzed. This model points a new direction for planners and other stakeholders to better monitoring the new mobility development. Also, private companies can benefit from this by avoiding unnecessary investment at an early stage.

This paper notes some limitations that need to be improved in a future study. i) Two reference cities are specially selected while only one city’s data is input into the model training. When considering the spatial heterogeneity among cities that can profoundly influence the accuracy of the prediction model, the same feature that affects the DSB’s spatial distribution can be different among cities. More cities with rich data in DSB distribution and features shall be introduced to the model to accumulate its transference ability. ii)The DSB is always considered as a primary transfer mode to connect other major transportation. The results are not noticeable to reveal this function through its feature importance ranking. The reason may be caused by only the midnight timepoint is considered while the DSB’s spatial distribution can be dynamic in both the short term and the long term. iii) As the GDBT model has shown its good prediction ability, the comparison with other deep learning methodologies is needed to prove its reliability in the future.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows. Study conception and design: D. Xiao, T. Gu, and I. Kim; data collection: D. Xiao, T. Gu, Y. Bao; analysis and interpretation of results: D. Xiao; draft manuscript preparation: D. Xiao, Y. Bao, and I. Kim. All authors reviewed the results and approved the final version of the manuscript.
REFERENCES


