Links between green space and public health
A bibliometric review of global research trends and future prospects from 1901 to 2019
Zhang, Jinguang; Yu, Zhaowu; Zhao, Bing; Sun, Ranhao; Vejre, Henrik

Published in:
Environmental Research Letters

DOI:
10.1088/1748-9326/ab7f64

Publication date:
2020

Document version
Publisher’s PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Links between green space and public health: a bibliometric review of global research trends and future prospects from 1901 to 2019

To cite this article: Jinguang Zhang et al 2020 Environ. Res. Lett. 15 063001

View the article online for updates and enhancements.
Links between green space and public health: a bibliometric review of global research trends and future prospects from 1901 to 2019

Jinguang Zhang¹,², Zhaowu Yu²,³,⁴, Bing Zhao¹, Ranhao Sun³ and Henrik Vejre²

¹ College of Landscape Architecture, Nanjing Forestry University, Nanjing 210037, People’s Republic of China
² Department of Geosciences and Natural Resource Management, Faculty of Science, University of Copenhagen, Copenhagen 1958, Denmark
³ State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, People’s Republic of China
⁴ Shanghai Key Lab for Urban Ecological Processes and Eco-Restoration, Shanghai, People’s Republic of China

E-mail: zhyu@ign.ku.dk

Keywords: bibliometric review, green spaces, physical health, mental health, ecosystem health, mechanism and pathway, health-based planning

Abstract

Considerable specific cross-sectional and review studies have linked exposure to green spaces to improving public health, but there is no bibliometric review attempting to systemically and retrospectively analyze these existing articles. Here we aim to uncover global research status, trends, and future prospects in green spaces and health (G-H) research then propose a framework for the underlying mechanisms and pathways that link green space to public health. We obtained 18 961 G-H research publications from the core Web of Science collection from 1901 to 2019, analyzing the characteristics of publication outputs, key scientific disciplines, and differences in performance between countries and institutions. Besides, content analysis was conducted to analyze the keywords, including keyword co-occurrence analysis and keyword clustering analysis. We found: (1) a steady quantitative increase in publications, active journals, and involved countries and institutions since the 1990s; (2) a significant increase and changes in G-H related interdisciplinary categories, with environment-related disciplines becoming the mainstream; (3) research focus and trends that were identified based on the analysis of high-frequency co-occurring keywords; (4) three main knowledge domains, namely, green spaces and physical health, mental health, and ecosystem health, that were identified and visualized based on keyword clustering analysis; (5) a framework of underlying mechanisms and pathways linking green space to public health that is proposed based on visualization of the three main knowledge domains. We suggest that the main challenge of G-H research is to further clarify in-depth the underlying mechanisms and pathways from multiple perspectives, including multiple nations, disciplines, and study designs. The lack of co-occurring keywords and clustering information related to social well-being suggests that research related to ‘social health’ is lacking. Based on a clear understanding of the quantity, quality, and characteristics of green space for public health, a health-based environmental plan should be proposed in the future.

1. Introduction

Rapid urbanization has dramatically changed the way people live over the past several decades (Un 2015). It has also reduced the possibility of urban residents coming into contact with green spaces (Mackerron and Mourato 2013, Bratman et al 2015, Ekkel and de Vries 2017). As a result, many studies have explored the negative effect of reduced contact with green space upon public health and the health effects of exposure to green spaces (Hartig et al 2014, Dadvand et al 2015, Gascon et al 2016). In general, exposure to green spaces can have a positive effect on public health, including overall health
(Dadvand et al 2016); birth weight (Hystad et al 2014, Markevych et al 2014); stress, depression, and anxiety (Thompson et al 2012; Mceachan et al 2016, Gascon et al 2018, Klompmaker et al 2019); hyperactivity disorder (Amoly et al 2014); postoperative rehabilitation (Ulrich 1984); obesity and type II diabetes (Amoly et al 2014, Ghimire et al 2017); various infectious diseases (Wong and Jim 2016); cardiovascular disease (Shen and Lung 2016); respiratory disease (Vienneau et al 2017); and all-cause mortality (Van den Berg et al 2015). However, green spaces occasionally have negative effects on public health, such as allergic diseases (Cariñanos and Casares-Porcel 2011), infections (Löhmus and Balbus 2015), injury (2015 Bortolini et al 2016), and even crime (Kimpton et al 2017). The reason for these ‘conflicting’ results is mainly due to a lack of clarity on the mechanisms underlying linkages between green spaces and public health (Markevych et al 2017). In addition, over-reliance on cross-sectional studies, different definitions of green space, and the different characteristics (e.g. size, type, normalized difference vegetation index, leaf area index, proximity, and accessibility) of green space, also cause uncertainty (van den Bosch and Sang 2017, Lai et al 2019). Therefore, a more process-oriented review of relevant research is needed to deepen our understanding of the complexities of the underlying pathways.

Existing reviews and meta-analyses primarily consider the link between green space and public health and collect evidence of health effects (Mcmahan and Estes 2015, Gascon et al 2016, 2018). Only a few specific reviews consider potential mechanisms and pathways that link green space (or nature) to public health. For instance, Hartig et al (2014) proposed four possible pathways between nature and health, including air quality, physical activity, social cohesion, and stress reduction. Among them, physical activities are the most prominent, which is the main focus of this review. Kuo (2015) identified 21 pathways empirically tied to nature and that have implications for specific physical and mental health outcomes; these pathways include environmental factors (e.g. filtering of pollutants from the air), physiological and psychological states (e.g. enhancing the immune system), and behaviors or conditions (e.g. promoting physical activity). In addition, Kuo also proposed criteria for identifying a possible central pathway, that is, enhanced immune functioning. A recent review by Markevych et al (2017) proposed potential underlying biopsychosocial pathways from a transdisciplinary standpoint based on the Expert Workshop, identifying three domain areas of green space, which are reducing harm (e.g. reducing exposure to environmental stressors), restoring capacities (e.g. attention restoration and physiological stress recovery), and building capacities (e.g. promoting physical activity and facilitating social cohesion). Additionally, the pathways linking green spaces to specific health effects were also highlighted. (de Jesus Crespo and Fulford 2018) conducted causal criteria analysis to characterize the ecosystem health (Eco-health) literature, reviewing the evidence for the full pathways between ecosystem (green space), ecosystem processes (e.g. ecosystem services), and health outcomes. They found sufficient evidence to suggest the link between ecosystem services of green spaces to lower risk of gastrointestinal disease and heat morbidities, where the ‘buffering’ ecosystem services can clean water, and mitigate water and heat hazards (Gao et al 2019). (Bratman et al 2019) proposed a pathway that would leverage existing knowledge about the impact of green space on mental health and ultimately incorporate it into ecosystem services assessments, involving four steps: nature features (size, type, and quality), exposure (e.g. proximity to nature, time spent in contact with nature), experience (interaction, doze), and effects (mental health).

Although burgeoning literature on green space and health (G-H) research may have revealed many potential mechanisms, existing systematic reviews or meta-analyses do not describe a consistent and persuasive linkage between green space and public health. The reasons for this include the limited review literature, neglect of interdisciplinary research, and some meta-analyses which have high levels of heterogeneity and include some poor-quality studies (Hunter et al 2019, Lai et al 2019). It is clear that G-H studies related to various potential mechanisms involve multiple disciplines and categories and hence should be comprehensively analyzed to understand developments, determine research status, and highlight cutting-edge trends.

Bibliometric reviews, a common research instrument for comprehensive analysis, have been used to explore and visualize scientific progress in many health-related fields recently, such as rural environment and health (Zhang et al 2019), childhood obesity (Gehanno et al 2019), nursing research (Dardas et al 2019), emerging technology in cancers (Wang et al 2019), and healthcare informatics (Gu et al 2017, Kokol et al 2018). To date, according to the search results of Web of Science (WoS) and Google Scholar, research on bibliometrics and visualization methods has not been used to dig deeper into and reveal the full panorama of, G-H research. In addition, co-occurrence keyword analysis of bibliometric reviews presents a comprehensive research tool that, in conjunction with a specific systematic review framework (appendix A (stacks.iop.org/ERL/15/063001/mmedia)), may help to explore the underlying mechanisms of G-H research to some extent (Hartig et al 2014, Kuo 2015, Markevych et al 2017; de Jesus Crespo et al 2018, Bratman et al 2019).

In this study, we adopted bibliometric analysis and newly developed visualization tools (e.g. VOS
viewer) to map the global research status and frontier trends of G-H research from multiple perspectives, providing researchers and practitioners with comprehensive and in-depth understanding of the salient research themes, as well as the potential pathways framework linking green space to public health. Four specific objectives of this study are: (1) to reveal the overall development status of G-H research by analyzing publication numbers, active journals, and research shifts within WoS categories; (2) to analyze patterns in publication performance for G-H research around the world over the past few decades by countries and institution; (3) to identify research focus and frontier trends of the current topic distribution in the field; and (4) to propose a framework of potential pathways linking green space to public health.

2. Method and data

2.1. Analysis methods

The bibliometric analysis describes the characteristics of specific categories of literature through various mathematical and statistical techniques, assesses the performance and cooperation of authors, institutions, and countries, discovers research priorities, and reveals research trends and future prospects. Here, we adopted this approach to quantify green space and health (G-H) research from 1901 to 2019. Typical techniques of bibliometric analysis include co-authorship analysis, co-occurrence analysis, co-citation analysis, and coupling analysis. Regarding the tools, CiteSpace 5.5.R2 (Chen 2006) and VOS viewer 1.6.12 (Van Eck and Waltman 2009) were selected to conduct the bibliometric analysis for this study. Besides, ArcGIS 10.5 software was used to generate geographic visualization maps.

Additionally, all keywords (including author keywords and keyword plus) extracted from a large number of pieces of literature by CiteSpace software were conducted to content analysis (Si et al 2019). There were two main steps employed in the content analysis of keywords to provide critical reviews. First, a high-frequency keyword co-occurrence analysis was conducted to indicate the research hotspots in a specific field (Chen et al 2014). Second, author keyword clustering analysis was applied to re-allocate retrieved information to investigate the evolution of a specific research area and provide reasonable predictions of future trends (Bouguettaya and Le Viet 1998). Notably, although keywords plus is as effective as author keywords in terms of bibliometric analysis for investigating the crucial knowledge of scientific fields, we only considered author keywords when generating structure (cluster) maps because the terms used for keywords plus are more broadly descriptive and less comprehensive in representing an article’s content (Zhang et al 2016).

2.2. Data acquisition and processing

Bibliographic data were retrieved from the core collection of Web of Science (WoS), initially with no time restriction on publication. A topic search of the WoS database was conducted. Phrases related to ‘green space’ and ‘health’ were considered as search terms, as presented in figure 1, and were used to collect all literature that contained at least one phrase related to each of these two terms in their titles, abstracts, or keywords. From these search terms, a total of 30,319 English documents were retrieved.
Due to the high proportion of articles as document type, and our research goals, this study only contains articles and review articles. In order to avoid interference from unrelated documents, some irrelevant categories were deliberately excluded, such as zoology and veterinary sciences (appendix B). Finally, a total of 18,961 documents in English (published before 1 October 2019) were finally retrieved. According to the results of the WoS search, the earliest paper on green space and public health was published in 1901, and so this study selected papers published from 1901 to 2019 for analysis. Subsequently, this period was subdivided into four-time intervals based on the publication of the literature (pre-1990s, 1990–1999, 2000–2009, 2010–2019) to more clearly explore the past research shifts and future trends.

3. Results

3.1. Quantitative analysis of research performance

3.1.1. Evolution of publication activity

The publication output for G-H research from 1901 to 2019 is summarized in figure 2. The first article in this field appeared in 1901, but there were only 11 documents retrieved from 1901 to 1989, which indicate a stagnation trend and very little development prior to the 1990s (which is the reason we did not consider this period in depth in the next analysis). After 1990, the annual number of publications started to rise steadily and has significantly increased since 2000, reaching its highest peak, 2,372 publications, in 2018. According to this growing trend, it could be predicted that the number of G-H research papers in 2019 will exceed the number of publications in 2018.

The number of active journals covering G-H research has also increased significantly over time, from only one in 1901 to the highest of 943 in 2018. In addition to the fact that the G-H study has attracted more attention from researchers in a variety of fields, the general increase of new scientific journals is probably also a major reason why there were more journals publishing studies on public health. There were only a small number of journals, such as *Environmental Health Perspectives*, which published G-H related studies in the pre-1990s. In the 1990s and 2000s, the journal with the largest number of articles is *Social Science & Medicine*. In the 2010s, *PLoS One*, became the journal with the largest number of publications (table 1).

3.1.2. Research shifts by Web of Science category

According to WoS’s classification system for subjects, a paper may belong to more than one subject area. The latest version of WoS classifies journals into 249 subject categories. Between 1901 and 2019, G-H research has been associated with approximately 50% of all subject categories, which suggests it has been involved in a very broad range of research areas. In the pre-1990s, early studies mainly described the role of parks in promoting public health. After 1990, the number of WoS categories has gradually increased. Figure 3 reveals the significant structural changes in Web of Science categories for G-H studies. During the 1990s, ecology ranked first in terms of the number of publications, followed by environmental science. Since the 2000s, environmental science has ranked the highest on the list of WoS categories for G-H research and has occupied a gradually larger proportion across the 2010s. In addition to environmental science, other categories related to the environmental field (environmental research) have developed rapidly in the last three decades. The
presence of many categories related to the environment indicates that environmental impacts are closely related to the processes linking green spaces to public health. Urban studies, multidisciplinary science, and geography newly appeared on the top 10 list of subject categories associated with publications in the 2010s. Moreover, urban studies became the fifth most frequently occurring subject category in the 2010s,
Table 2. The top 10 most productive countries ranked in descending order over three decades.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GDP (%)</td>
<td>Population (%)</td>
<td>Publications</td>
</tr>
<tr>
<td>USA</td>
<td>26.72</td>
<td>4.62</td>
<td>538</td>
</tr>
<tr>
<td>UK</td>
<td>4.68</td>
<td>0.97</td>
<td>125</td>
</tr>
<tr>
<td>Australia</td>
<td>1.26</td>
<td>0.31</td>
<td>78</td>
</tr>
<tr>
<td>Canada</td>
<td>2.17</td>
<td>0.50</td>
<td>66</td>
</tr>
<tr>
<td>Germany</td>
<td>7.68</td>
<td>1.36</td>
<td>55</td>
</tr>
<tr>
<td>France</td>
<td>5.05</td>
<td>1.00</td>
<td>32</td>
</tr>
<tr>
<td>Italy</td>
<td>4.29</td>
<td>0.94</td>
<td>29</td>
</tr>
<tr>
<td>Netherlands</td>
<td>1.39</td>
<td>0.26</td>
<td>28</td>
</tr>
<tr>
<td>Spain</td>
<td>2.08</td>
<td>0.67</td>
<td>27</td>
</tr>
<tr>
<td>Japan</td>
<td>15.3</td>
<td>2.10</td>
<td>17</td>
</tr>
</tbody>
</table>

GDP(%) represents the cumulative gross domestic product(GDP) of each country in every decade as a percentage of the global cumulative GDP; Population (%) represents the proportion of the population in the last year of every decade to the total world population in that year.
which suggests, to a certain degree, that health factors were increasingly considered in urban planning and design (Rydin et al 2012, Wolch et al 2014).

3.1.3. Publication performance: countries and institutions

With respect to publication performance by country, almost all the papers that were published in the pre-1990s came from the USA. Moreover, the USA has always been the leader in G-H research, contributing the most publications in the other three time periods and accounting for 42.3% (1990s), 34.7% (2000s), and 22.8% (2010s) of total publications respectively. Table 2 shows the top 10 countries (number of publications) that participated in G-H research from the 1990s onward. The top 10 countries in terms of numbers of publications contributed 76.3% (1990s), 68.9% (2000s), and 61.6% (2010s) of total publications, respectively, during the 30-year period. In addition, the top 10 countries in the three stages are mostly developed countries, which account for 68.42%, 59.40%, and 57.44% of the world’s gross domestic product (GDP) respectively in the three decades. This observation indicates that the stronger a country’s economic capacity, the richer its research in the green spaces and health field is likely to be. However, the total population of these countries accounts for only 12.75%, 29.59%, and 27.95% of the world’s population, respectively. It also suggests that developing countries with large populations, especially low-income groups, which are the most vulnerable groups to health issues, have insufficient theoretical and practical results regarding the G-H research. G-H research has developed rapidly in China, and the volume of publications from China has grown from only 14 in the 1990s to 1422 in the 2010s.

Institutions in the USA were always the most productive in the pre-1990s period. However, they do not always rank in the leading positions in the other three decades. Table 3 lists the top 10 most productive institutions. In these three periods, US institutions only ranked first in the 2000s whilst in the other two periods, their positions were taken by Australia and China. In China, the institutional rankings of the Chinese Academy of Sciences rose sharply from the 2000s to the 2010s and became the most productive institutions in the 2010s.

Geographic maps were applied to indicate the spatial distribution of institutional publication activity, which refers to the frequency by which an institution was extracted from the author’s affiliations. According to the definition proposed by (Wu and Ren 2018)), low, middle, and high activity correspond to \( \leq 5 \), 5–10, and \( \geq 10 \) publications for each period, respectively. In the pre-1990s, G-H research institutions were few and mainly distributed in big cities of the USA, such as New York (this period is not shown on the geographic maps). In the 1990s, only a few US institutions showed high publishing activity, some institutions in Europe showed a middle level of publishing activity, and almost all other institutions showed low publishing activity levels. In the 2000s, institutions involved in G-H research increased rapidly and publication activity reached a high level. However, high G-H research publication
activity was mainly restricted to the USA, Europe, and Australia. There were only several high-activity institutions in East Asia. In the 2010s, G-H research activity expanded significantly in East Asia, and G-H research institutions had now expanded to six continents (figure 4).

3.2. Research focus, trends, and prospects

3.2.1. Research focus and trends: keyword analysis

Analyzing keywords can uncover the most interesting areas or issues in specific research fields. Unfortunately, documents retrieved from WoS in the pre-1990s period were excluded, as their keywords are not available. Here, 1737, 10 294, and 52 307 keywords were obtained for the 1990s, 2000s, and 2010s, respectively (appendix C). The dramatic increase in the number of keywords indicates a broadening emphasis of G-H research. Duplicate detection and merging were performed in Excel, followed by a manual check. The frequency of co-occurrence of keywords can reflect, to a certain extent, hotspots within a research theme, the methods, and
The 20 most frequently co-occurring keywords ranked in descending order over three decades.

<table>
<thead>
<tr>
<th>Keyword(s)</th>
<th>1990–1999 Frequency</th>
<th>Keyword(s)</th>
<th>2000–2009 Frequency</th>
<th>Keyword(s)</th>
<th>2010–2019 Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stress</td>
<td>72</td>
<td>Health</td>
<td>265</td>
<td>Health</td>
<td>2194</td>
</tr>
<tr>
<td>Growth</td>
<td>46</td>
<td>Stress</td>
<td>186</td>
<td>Physical activity</td>
<td>1288</td>
</tr>
<tr>
<td>Health</td>
<td>43</td>
<td>Community</td>
<td>148</td>
<td>Environment</td>
<td>1015</td>
</tr>
<tr>
<td>Mortality</td>
<td>42</td>
<td>Mortality</td>
<td>136</td>
<td>Impact</td>
<td>950</td>
</tr>
<tr>
<td>Plant</td>
<td>34</td>
<td>Environment</td>
<td>127</td>
<td>Climate change</td>
<td>812</td>
</tr>
<tr>
<td>Population</td>
<td>34</td>
<td>Management</td>
<td>113</td>
<td>Green space</td>
<td>780</td>
</tr>
<tr>
<td>Natural disaster</td>
<td>28</td>
<td>Growth</td>
<td>112</td>
<td>City</td>
<td>735</td>
</tr>
<tr>
<td>Children</td>
<td>27</td>
<td>National park</td>
<td>110</td>
<td>Stress</td>
<td>732</td>
</tr>
<tr>
<td>Model</td>
<td>27</td>
<td>Population</td>
<td>110</td>
<td>Management</td>
<td>598</td>
</tr>
<tr>
<td>Natural history</td>
<td>27</td>
<td>Conservation</td>
<td>108</td>
<td>Exposure</td>
<td>579</td>
</tr>
<tr>
<td>Community</td>
<td>26</td>
<td>Physical activity</td>
<td>103</td>
<td>Biodiversity</td>
<td>570</td>
</tr>
<tr>
<td>Symptom</td>
<td>26</td>
<td>United States</td>
<td>102</td>
<td>Ecosystem service</td>
<td>559</td>
</tr>
<tr>
<td>Photosynthesis</td>
<td>25</td>
<td>Forest</td>
<td>102</td>
<td>Air pollution</td>
<td>543</td>
</tr>
<tr>
<td>Depression</td>
<td>23</td>
<td>Children</td>
<td>102</td>
<td>Community</td>
<td>526</td>
</tr>
<tr>
<td>California</td>
<td>22</td>
<td>Model</td>
<td>98</td>
<td>Park</td>
<td>525</td>
</tr>
<tr>
<td>Dynamics</td>
<td>22</td>
<td>Plant</td>
<td>96</td>
<td>Built environment</td>
<td>523</td>
</tr>
<tr>
<td>Experience</td>
<td>20</td>
<td>Impact</td>
<td>94</td>
<td>Children</td>
<td>513</td>
</tr>
<tr>
<td>Behavior</td>
<td>19</td>
<td>Risk</td>
<td>86</td>
<td>Mortality</td>
<td>512</td>
</tr>
<tr>
<td>Leave</td>
<td>19</td>
<td>Vegetation</td>
<td>83</td>
<td>Association</td>
<td>509</td>
</tr>
<tr>
<td>Soil</td>
<td>19</td>
<td>Women</td>
<td>80</td>
<td>Mental health</td>
<td>500</td>
</tr>
</tbody>
</table>

the content of retrieved documents. High-frequency keywords that appear in all three decades suggest that the research terms represented by those words remain a continued focus. Changes in high-frequency words indicate a shift in research emphases, representing the trends of the G-H research to some extent.

The relationship between green space and stress has been a major focus of attention during recent decades (table 4). Stress is the process by which an individual responds psychologically, physiologically, and often with behaviors, to a situation that challenges or threatens their well-being (Baum et al. 1985). The vast majority of research to date has suggested that exposure to green space can relieve the pressure that residents feel and help with recovery from stress (Ulrich et al. 1991). Another topic that has been a hotspot and continues to be important is the association between green space and mortality. For example, (Gascon et al. 2016) reviewed the available evidence on the relationship between long-term exposure to residential green spaces and mortality in adults, and found support for the hypothesis that living in areas with more green space reduces mortality, mainly cardiovascular disease mortality. Additionally, vulnerable groups, in particular, children have been considered as the main target in G-H research during the three decades. For instance, (Christian et al. 2015) reported a positive relationship between neighborhood green spaces and early child health development and related behaviors. Other researchers have found that green spaces are crucial for young children’s mental health (Engemann et al. 2019), attention (Faber Taylor and Kuo 2009), and motor development (Fjørtoft and Sageie 2000).

Regarding newly appearing high-frequency words, ‘physical activity,’ which first became a high-frequency word (ranked 11th) in the 2000s, has now dramatically risen to be ranked in the top 2 in the 2010s. Physical activity, which has been considered as a critical pathway for linking nature to human health, promotes physical and mental health across the life span (Hunter et al. 2015). Other emerging high-frequency words mainly originate from environmental science, such as ‘climate change,’ ‘biodiversity,’ ‘ecosystem services,’ and ‘air pollution.’ Due to rapid population growth and urbanization, environmental problems that are not conducive to health, such as heat hazards, air pollution, and climate change, received the highest attention. Ecosystem services afforded by green spaces as nature-based solutions for improving public health (e.g. improving the above-mentioned environmental problems) was also becoming a significant focus since the 2010s (van den Bosch and Sang 2017, Barnes et al. 2018).

3.2.2. Main knowledge domains: keyword-clustering
To identify the main knowledge domains, VOSviewer that has unique advantages in mapping knowledge domain label structure (Van Eck and Waltman 2009) was adopted to indicate the co-occurrence relationships of keywords, which can clearly indicate the internal composition or structure of a field. A total of 38 637 author keywords appeared in the retrieved documents were identified by VOSviewer (appendix D). Furthermore, keywords with a frequency of more than 25 appearances were collected, and small clusters (less than 30) were merged acquiescently. Finally, a total of 284 qualified keywords were successfully identified and classified into three groups (blue, green, red). As shown in figure 5, each circle represents a keyword, and the size of the circle
denotes the frequency of that keyword. Lines connecting the circles indicate relationships between the two keywords, and colors identify the different clusters, which correspond to the main specific areas of G-H research. Hence, based on the clustering group and the important keyword (high-frequency keyword), we can clearly identify the research main domains and focus.

Furthermore, in order to determine the specific name of each domain, we referred to the definition of health by World Health Organization (WHO), that is, ‘health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity’ (World Health 1948). Conceived in this way, health encompasses physical health, mental health, social health (well-being), and the absence of disease. Additionally, green spaces provide a wide range of ecosystem services (ESs), in particular, regulating services (e.g. reduction of heat) (Yu et al 2019, 2020a, 2020b, Yang et al 2020), and supporting services (e.g. maintenance of biodiversity) (Kong et al 2010), many of which are positively linked to reducing risk of disease (e.g. heat morbidity) (Hartig et al 2014, Markeyvych et al 2017, Bratman et al 2019). Hence, an increasing number of scholars are considering ecosystem health (eco-health) (Su et al 2010, O’Brien et al 2016, de Jesus Crespo and Fulford 2018).

Combining these concepts with the keyword clustering analysis, we could finally determine the specific domains in G-H research. Due to the lack of ‘social health (well-being)’ related keyword, we deliberately proposed and defined the three clustering groups as green spaces and physical health (G-PH), green spaces and mental health (G-MH), and green spaces and ecosystem health (G-EH). In addition, according to the number of clustering terms (appendix E), we can roughly estimate that the G-EH studies occupy the largest proportion of retrieved literature, accounting for about 48.8%, followed by the G-MH related literature, accounting for about 30.3%, and G-PH researches, accounting for about 20.9%.

3.2.3 Cluster 1 (blue): green spaces and physical health (G-PH)
In the blue cluster, the keywords ‘health,’ ‘public health,’ ‘well-being,’ ‘green space,’ and ‘parks’ are the nodes of research term that are highlighted. In addition, cultural ecosystem services provided by green spaces also presented in this cluster, such as the nodes ‘leisure’ and ‘recreation.’ However, the biggest node in this cluster is ‘physical activity,’ strongly linking with above-mentioned research terms, indicating that green spaces (e.g. parks) promote certain types of physical activity (‘exercise,’ ‘walking,’ and ‘cycling’), and thus has a significant contribution to physical health (e.g. ‘obesity’). Notably, there is some evidence that nature-based physical activity (green exercise) achieves more physical health benefits than equivalent exertion in indoor or constructed settings (Thompson Coon et al 2011, Barton et al 2016, Frumkin et al 2017).

3.2.4 Cluster 2 (green): green spaces and mental health (G-MH)
The second research cluster refers to the linkage between green space and mental health. Mental health is defined by the WHO as ‘a state of well-being in which an individual realizes his or her own potential, can cope with the normal stresses of life…’. With regard to this conceptualization, mental health encompasses the presence of psychological well-being and the absence of mental illness. The largest circle, ‘mental health,’ mainly links to ‘stress,’ ‘depression,’ and some disease-related keywords, such as ‘mortality,’ ‘epidemiology,’ and ‘hypertension.’ Many studies have demonstrated that exposure to green space is positively associated with mental health, including mental illness and the presence of psychological well-being (Bratman et al 2019).

3.2.5 Cluster 3 (red): green spaces and ecosystem health (G-EH)
The red cluster is the largest of the three clusters, and corresponds to the link between green space and ecosystem health, with a focus on environmentally relevant hotspots. The largest node, ‘ecosystem services,’ mainly links to ‘climate change,’ ‘air pollution,’ ‘green infrastructure,’ ‘drought,’ ‘urban heat island,’ ‘sustainability,’ ‘conservation,’ and ‘biodiversity.’ Green space plays a crucial role in advancing the ecological sustainability of cities and promoting human health by providing a wide range of ecosystem services, such as water cleansing, mitigation of water hazards, improving air quality, reducing noise, maintaining biodiversity, and preventing heat stress.

3.3 Linkage between green space and public health: What are the pathways?
Although cluster analysis can group related terms together, the nodes are not ordered and cannot clearly indicate the underlying connections between green space and public health. Hence, we filtered out relevant high-frequency keywords and selected specific review literature (appendix A) in the retrieved G-H documents database to explore the potential pathways linking green space to public health. The underlying mechanism linking green space and public health is very complex. Here we mainly explore the pathways linking green space to health in three main knowledge domains. It is worth noting that exposure to green spaces is a prerequisite for achieving the health benefits of green spaces, especially for physical and mental health. Exposure is a broad term referring to a variable amount or type of exposure to green space and can be evaluated by multiple indicators, such as the availability of green space (surrounding
Figure 5. The author keyword cluster analysis: physical health (blue), mental health (green), and eco-health (red).

3.3.1 Pathway 1: linking green space to physical health

Living near (or having easily accessible) green spaces, with safe and attractive areas for activity, promotes physical health: the underlying mechanism of this effect is physical activity levels (Pretty et al. 2003). A large body of literature has confirmed that physical activity in green spaces (green sports) produces greater health benefits than sports activities in the built environment (Triguero-Mas et al. 2015, Dadvand et al. 2016, van den Berg et al. 2019). Over three broad activity domains (work, active transport, and leisure), green spaces are considered essential for promoting physical activity via active transport (walking or cycling) and leisure (sport or recreation) (Hartig et al. 2014). For example, higher levels of greenness in the environment may lead people to favor walking or cycling over other transport modes by making routes to the destinations more attractive, although other factors such as distance, infrastructure, and safety are also important (Heinen et al. 2010). Nevertheless, there are potential negative effects associated with physical activity (e.g. sports injuries). Green space can offer opportunities for most beneficial ‘green exercises,’ which are positively associated with a lower risk of chronic diseases such as obesity, type II diabetes, and hypertension. Therefore, physical activity is a crucial intermediate factor in establishing the links between green space and public health, especially for physical health (Triguero-Mas et al. 2015, Hunter et al. 2019).

3.3.2 Pathway 2: linking green space to mental health

Exposure to green spaces can have a positive impact on reported psychological well-being. Although the mechanisms underlying these links remain unclear, perceptions and experiences that are mainly immaterially related to green space, are probably important mechanistic connections that promote better mental health. Knowing and viewing are two specific ways people perceive green spaces (Russell et al. 2013). Knowledge of belonging to a community rather than being alone, derived through experiences of nature, is argued to be a plausible reason for the broadly positive psychological benefits of nature (Mayer et al. 2009). Views of green spaces have been repeatedly associated with health benefits, especially for reducing stress levels, such as perceived stress, job stress, and driving stress (Parsons et al. 1998, Bringslimark et al. 2011). Another mechanistic channel experiences, which represents an interaction with green spaces (Bratman et al. 2012). To date, around 150 human-nature interaction patterns have been generated and cataloged (Bratman et al. 2019). Long-term interaction with green spaces has been extensively documented to have a positive effect on mental health, suggesting improved happiness and subjective well-being. For example, experiencing green spaces promotes restoration (e.g. increased attention, and reduced stress, fatigue, and irritability) better than does other experiences that have been studied (such as
interacting with friends) (Ottosson and Grahn 2005). In addition, exposure to green space has been associated with reductions in mental illness, due to the benefits that are perceived or experienced regarding green spaces in terms of improving sleep, relieving stress, reducing rumination (a maladaptive pattern of self-referential thought), and subgenual prefrontal cortex activation, which are the major risk factors for mental illness, especially depression (Bratman et al 2015).

3.3.3 Pathway 3: linking green space to ecosystem health
Many studies have shown that there is an intermediate link between the existence of ecosystems and public health through buffering ecosystem services (ESs), which refers to the ability of ecosystems to ‘buffer’ health effects (Bratman et al 2012, Jackson et al 2013, Guerry et al 2015, de Jesus Crespo and Fulford 2018). The processes of buffering ESs that are most relevant to public health can be broadly classified into the following categories: clean air and water; noise reduction; flood mitigation; heat effect reduction; climate stability; biodiversity conservation, etc. Therefore, green space can modify health outcomes to a certain extent and reduce the risk of clinical diseases such as gastrointestinal diseases, respiratory diseases, and cardiovascular diseases, through the causal pathway of buffering ESs. For example, green space ‘water cleansing’ can help prevent the gastrointestinal disease from drinking or recreational use of infectious or toxic water (Katukiza et al 2014). Also, green space ‘mitigation of heat effects’ can help avoid extreme temperatures through shade and evapotranspiration, lowering heat morbidities during heatwaves (Bouchama and Knochel 2002). Another focus for plausible causal pathways of buffering ESs (supporting services) and important concerns, is the role of greenspace biodiversity in promoting eco-health. One explanation is that exposure to diverse microbiomes in green spaces (e.g. soils, plants, water) helps train the immune system to accurately distinguish dangerous from helpful bacteria and modulates immune function, which can affect a number of health outcomes, such as in the areas of cancer, allergies, and autoimmune diseases (Hanski et al 2012, Ruokolainen et al 2015, Von Hertzen et al 2015).

In addition to the causal mechanisms and pathways of the three domains mentioned above, some extra explanations need to be outlined about the proposed framework. First, although we distinguished clusters of causal pathways with different colors (figure 6), the actual pathways are complexly intertwined, and multiple pathways are likely to be engaged simultaneously. For example, ecosystem services (cultural services) can also provide better aesthetic environments and recreational opportunities to promote physical activity, which is positively associated with physiological and psychological health. In addition, the strongest evidence suggests that the greatest benefits of physical activity are in the prevention and treatment of cardiovascular disease and related mortality as well as all-cause mortality (Thompson et al 2003, Van den Berg et al 2015). Second, all these beneficial health outcomes of green spaces may be affected by potential moderating and mediating factors, such as socio-economic status (SES), gender, age, preferences, occupation, culture, and individual perceptions (Triguero-Mas et al 2015, Akpinar et al 2016, Dadvand et al 2016, Sugiyama et al 2016). Third, the social benefits of interacting with green spaces (e.g. increased social cohesion), which can be regarded as ‘social health’ with respect to neighborhoods and communities, have received less attention than the other three health categories (Keniger et al 2013, Hunter et al 2019). Fourth, nature-based planning and health-based design should be considered when designers and planners advance the programmatic and policy goals for green spaces in order to achieve increased health benefits (Barnes et al 2018).

4. Discussion

4.1. Challenges: exploring potential mechanisms in-depth from multiple perspectives
This study has proposed a pathway framework based on three thematic domains generated by keyword clustering analysis. However, it is not thorough and comprehensive because it relies too much on keyword content analysis and insufficient in-depth reading of specific reviews. Hence, future research needs to further clarify potential mechanisms and pathways from multiple perspectives (including multiple countries, disciplines, and research designs).

Based on bibliometric methods to quantitatively analyze the performance of publications, this study found that the G-H research literature was mostly from highly developed countries, especially the United States and the United Kingdom, which is consistent with recent systematic reviews (Markeych et al 2017, Hunter et al 2019, Lai et al 2019). However, there are few studies in Africa, Asia, and South America, which may limit the generalizability of findings for green space health effects. Therefore, research studies in these developing regions are needed, as they are currently experiencing a rapid urbanization process. For example, exploring the potential pathways by which green spaces can buffer adverse effects of urbanization in sub-Saharan African regions could achieve considerable health returns, because these regions will be experiencing the highest rates of urbanization and population growth over the next several decades (Un 2015).

Another challenge is that G-H research involves approximately 50% of all subject categories. As we have noted, cross-disciplinary research is hence necessary to explore biopsychosocial pathways and
4.2. Gaps: lack of social health research

As defined by the WHO, health includes physical health, mental health, social health, and the absence of disease and infirmity. According to the high-frequency keyword analysis and keyword cluster analysis in the previous sections, there are many studies that consider physical and mental health, as well as the relationship between green space and lowered risk of diseases. Although it is widely recognized that social contact with green spaces can influence several health outcomes (Rios et al. 2011, Fone et al. 2014), the absence of co-occurring keywords and clustering information related to social wellbeing, indicates a lack of social health-related research. Our results confirmed this opinion from a scientific and quantitative perspective, which is in line with previous studies. For example, a systematic review by (Hartig et al. 2014) found that the environmental correlates of social cohesion have received little research attention. Similarly, (Keniger et al. 2013) suggested that the social benefits of interacting with nature have received less attention than physical health, cognitive performance, and psychological wellbeing, despite the potential for significant consequences arising from the former. With respect to the underlying mechanisms linking green spaces to health, social cohesion may be an important mediating factor, as it is positively related to people’s feelings of loneliness and shortages of social support (Maas et al. 2009). However, this has not been widely observed in scientific research (Triguero-Mas et al. 2015), and more thorough and extensive studies are needed to fill this gap.

4.3. Directions: health-based environment planning

Existing literature related to G-H research mainly considers two factors, that are: (i) to explore whether or to what extent health (physical and mental health, psychological wellbeing, and other health outcomes) is positively associated with green spaces (Bowler et al. 2010, Mcmahan and Estes 2015, Gascon et al. 2016); or (ii) to primarily explain the underlying pathways and mechanisms linking green spaces to
public health (Markevych et al. 2017). However, in addition to continually exploring the underlying mechanisms in-depth, the next step should be to focus on the processes that produce health benefits and how they can be best promoted. Decision-makers and urban planners who aim to improve public health in built environments should be provided the guidance of health-based environment planning to best promote human health, regarding landscape composition (e.g. plants species, water, and effective activity areas), dose-response relationships (e.g. duration spent in green spaces), accessibility (e.g. proximity to public parks), qualities (e.g. amenities, tree canopy coverage), and features (e.g. size, type) of green spaces (Barnes et al. 2018). For example, it is widely acknowledged that pollen emission by higher plants affects human health during the pollination period, which might cause an allergic response to humans. It is estimated that by 2025 more than 50% of all Europeans will suffer from at least one type of allergy, with no age, social or geographical distinction. It is estimated that by 2025 more than 50% of all Europeans will suffer from at least one type of allergy, with no age, social or geographical distinction (European Academy of Allergy and Clinical Immunology, https://www.eaaci.org/documents/EAACI_Advocacy_Manifesto.pdf). However, we found that the ‘allergy’ or related keywords are almost never detected in retrieved literature database, which indicates that rare attention has been paid to the actual or potential allergenicity of green spaces (e.g. plant species); in this regard, health-based environment planning guidance or the criteria used in selecting urban tree species should propose to avoid using the mass monospecific pollen (Álcazar et al. 2004), and oppose to the botanical sexism in dioecious species, which currently tend to select male plants that may release large amounts of pollen, while discriminated female plants due to the problems such as fruit fall, insects or bad smell (Cariñanos et al. 2014).

4.4. Limitations
Although this study objectively reveals the current status and trends of global G-H research, there were several limitations that should be discussed. First, this study only used publications from the WoS database and did not include other databases such as Scopus, Medline, and Zetoc, which may have limited the scope of our data collection. Second, we may have introduced errors in data formatting by relying on CiteSpace to conduct our analysis of keyword frequency. Third, the definition of a minimum threshold of keyword co-occurrences for analysis in CiteSpace may influence the identification keyword terms and the different parameter settings for VOSviewer may produce different clustering visualization results.

5. Conclusion
The number of papers exploring linkages between green spaces and public health has multiplied in recent years, mainly due to the rapid growth of urban populations and lifestyle-related diseases that contribute to the major causes of mortality. To our knowledge, this review is the first of its kind to perform a systematic review with a bibliometric method to quantitatively reveal the global research status and trends in G-H research based on 18 961 publications retrieved from the core WoS collection. Our results indicated a steady increase in publications, active journals, and the number of involved countries and institutions. In addition, a significant increase in changes in G-H related interdisciplinary categories were quantitatively revealed, with environment-related categories becoming the mainstream. Furthermore, keywords analysis, including co-occurrence analysis and cluster analysis, revealed the research focus, trends, and common themes. More importantly, we propose a framework for exploring potential mechanisms and pathways linking green space to public health, using knowledge domain visualization based on keyword analysis. Finally, we conclude that the study of G-H research is a growing line of interdisciplinary research worldwide.

This study retrospectively reviewed long-term trends of G-H research over the last century. The earliest literature on green space and health (G-H research) was retrieved in 1901. Interestingly, there were concerns about the negative health effects of the green spaces due to the mosquitoes in the newly constructed water park that might cause the spread of malaria. Then, especially in the recent three decades, people gradually realized the importance of green space for public health. To be specific, exposure to green space promoting mental health, in particular, relieving urbanites’ stress, were consistently research focus in three decades. Subsequently, ‘physical activity’ emerged and became a high-frequency word in the 2000s, accompanied by related- words such as exercise and walking. We thus speculated that scholars began to explore the physical health that afforded by green space through promoting physical activity. In the 2010s, mental health and physical health have been paid a lot of attention, and ecosystem health-related studies dramatically increased. In the future, more thorough and extensive studies are needed to consider associations between green spaces and social health. In addition, health-based environment planning research should be proposed in the future to help designers and practitioners achieve the ‘best’ improvements in public health.

Acknowledgments
This work was supported by the National Natural Science Foundation of China (Grant No. 41922007); Open Foundation of the State Key Laboratory of Urban and Regional Ecology of China (Grant No. SKLURE 2019-2-6); Shanghai Key Lab for Urban
Ecological Processes and Eco-Restoration (Grant No. SHUES2019A01); WEL Visiting Fellowship Program; and Chinese Scholarship Council (Grant No. 201908320579). We also thank two anonymous reviewers for their constructive comments.

Data availability

The data that support the findings of this study are available upon reasonable request from the authors.

References

Astell-Burt T et al 2014 Is neighborhood green space associated with a lower risk of type 2 diabetes? Evidence from 267,072 Australians Diabetes Care 37 197–201
Barnes M R et al 2018 Characterizing nature and participant experience in studies of nature exposure for positive mental health, an integrative review Front. Psychol. 9 2617
Bouguettaya A and Le Viet Q 1998 Data clustering analysis in a multidimensional space Inf. Sci. 112 267–95
Bowler D E et al 2010 A systematic review of evidence for the added benefits to health of exposure to natural environments BMC Public Health 10 439
Bringløm T et al 2011 Adaptation to windowlessness: do office workers compensate for a lack of visual access to the outdoors? Environ. Behav. 43 469–87
Christian H et al 2015 The influence of the neighborhood physical environment on early child health and development: a review and call for research Health Place 33 23–36
Dadvand P et al 2012 Surrounding greenness and pregnancy outcomes in four Spanish birth cohorts Environ. Health Perspect. 120 1481–7
de Jesus Crespo R and Fulford R 2018 Eco-Health linkages: assessing the role of ecosystem goods and services on human health using causal criteria analysis Int. J. Public Health 63 81–92
Engemann K et al 2019 Residential green space in childhood is associated with lower risk of psychiatric disorders from adolescence into adulthood Proc. Natl. Acad. Sci. 116 5188–93
Faber Taylor A and Kuo F E 2009 Children with attention deficits concentrate better after walk in the park J. Attent. Disord. 12 402–9
Fone D et al 2014 Effect of neighbourhood deprivation and social cohesion on mental health inequality: a multilevel population-based longitudinal study Psychol. Med. 44 2449–60
Frumkin H et al 2017 Nature contact and human health: a research agenda Environ. Health Perspect. 125 075001
Ghinnire R et al 2017 Green space and adult obesity in the United States Ecol. Econ. 136 201–12
Halonen J I et al 2014 Green and blue areas as predictors of overweight and obesity in an 8-year follow-up study Obesity 22 1910–1917.
Hunter R F et al 2013 The impact of interventions to promote physical activity in urban green space: a systematic review
Wu Z and Ren Y 2018 A bibliometric review of past trends and future prospects in urban heat island research from 1990 to 2017 *Environmental Review*. **27** 241–51
Yang G et al 2020 How can urban blue-green space be planned for climate adaption in high-latitude cities? A seasonal perspective *Sustainability*. **53** 101932
Yu Z et al 2020b Critical review on the cooling effect of urban blue-green space: a threshold-size perspective *Urban Forestry and Urban Greening*. **126630**