Research paper **Open Access**

Xiaorui Cao, Ruodan Lu*, Liang Guo and Jianya Liu*

Construction health and safety: A topic landscape study

DOI 10.2478/otmcj-2021-0027
Received: July 30, 2020; Accepted: July 14, 2021

Abstract: We aim to draw in-depth insights into the current literature in construction health and safety and provide perspectives for future research efforts. The existing literature on construction health and safety is not only diverse and rich in sight, but also complex and fragmented in structure. It is essential for the construction industry and research community to understand the overall development and existing challenges of construction health and safety to adapt to future new code of practice and challenges in this field. We mapped the topic landscape followed by identifying the salient development trajectories of this research area over time. We used the topic modeling algorithm to extract 10 distinct topics from 662 abstracts (filtered from a total of 895) of articles published between 1991 and 2020. In addition, we provided the most cited references and the most popular journal per topic as well. The results from a time series analysis suggested that the construction health and safety would maintain its popularity in the next 5 years. Research efforts would be devoted to the topics including "Physical health and disease", "Migrant and race", "Vocational ability and training", and "Smart devices." Among these topics, "Smart devices" would be the most promising one.

Keywords: construction health and safety, occupational health, topic landscape, content analysis, time series analysis

1 Introduction

Construction is a labor-intensive activity. According to the International Labour Organization (ILO), the world's estimated 111 million construction workers take around 15% of the total number of employed personnel (ILO, 2001). Health and safety are major concerns in construction jobsites. Construction workers are vulnerable to both work-related illness and injury. According to the Health and Safety Executive (HSE), within the UK, in 2019, there were 79,000 construction workers suffering from work-related ill health problems and 54,000 from non-fatal injuries every year (HSE, 2019). ILO estimates that there are at least 60,000 fatal accidents per year in worldwide construction jobsites, meaning that more than one fatal accident occurs every 10 min (ILO, 2003). Accidents and ill health have a financial cost. The UK construction industry, for instance, struggles with approximately £1.2 billion total cost and 2.0 million working days lost per year (HSE, 2019). The speed and strength with which Coronavirus disease 2019 (COVID-19) has struck the global construction industry is unprecedented. There is no doubt that all participants in the construction sector have and will continue to experience impacts on their operations in the post-COVID era. It is essential for employers and construction workers to improve the performance and prepare to adapt to a new world order and code of practice.

Occupational Safety and Health (OSH) of construction is generally a vast domain, interfacing and encompassing numerous disciplines and a large number of workplace and environmental factors. Finneran and Gibb (2013) claim that although the phrase and term OSH is used within the construction industry, the focus of safety and health is unevenly distributed. This is because researchers entering the OSH area carry with them specific culture and expertise which, in turn, creates a broad spectrum of research genres and stakeholders. This is also because construction itself is a complex, transient, and global sector, made up of various organizational sizes and affiliated companies. Health and safety can impact the whole life cycle of a project from planning, design, and construction to operationalization. Much research effort has been devoted to preventing, assessing, and reducing safety hazards as well as controlling risks on construction jobsites while improving workers' health and welfare and increasing working

***Corresponding author: Ruodan Lu,** Darwin College, University of Cambridge, UK. E-mail: rl508@cam.ac.uk
***Corresponding author: Jianya Liu,** Data Science institute, Shandong University, China
Xiaorui Cao and Liang Guo, Data Science institute, Shandong University, China

productivity. These research works observe and collect real-life construction workers' OSH data followed by analyzing them through qualitative and quantitative approaches. They span from workers' identity and living habits (Arcury et al. 2014) to workforce climate and evolution of safety culture (Siu et al. 2004; Arcury et al. 2012), from mental health and wellbeing (Kotera et al. 2019) to physical health and strain diseases (Antwi-Afari et al. 2017), from human factors/ergonomics (Cheng et al. 2013) and site activity management (Yi et al., 2016) to wearable and sensor-based devices (Lee et al. 2017; Jebelli et al. 2019).

Multidisciplinary knowledge in the OSH field is not only diverse and insightful, but also fragmented and multifaceted (Finneran and Gibb, 2013). There is a clearly pressing need for mapping of this ever-more complex landscape to help both the industry and research communities to conduct an efficient and effective literature review. This research is targeted at a wide range of audiences, including scholars, journal editors, and construction practitioners. It aims to provide an in-depth understanding of how OSH has evolved over time, distilling insights and shedding light on the points of consensus and divergences among scholars, while revealing research gaps and predicting future trends in the intellectual structure of the field. This goal is achieved by answering the following research topics: (1) How to effectively review the existing OSH literature and (2) how to reveal its hidden structure of distinct topics.

This research comprises a computer-based overview of the literature of OSH in construction over the period of 1991–2020. We mapped out the topic landscape followed by attempting to anticipate popular topics that will generally be found in seminal research in the future. To the best of our knowledge, this is so far the first systematic review of OSH in construction of this kind that covers many disciplines over decades, and the first attempt to forecast topic prevalence in this domain. The first contribution of this research lies in the fact that it uncovers a hidden structure of 10 distinct topics and development trajectories in a corpus comprising the abstracts of 662 most representative scholarly articles (filtered from a total of 895 abstracts) published in the past three decades. The corpus contains a broad range of author backgrounds, disciplinary influences, and research focuses. It enables readers to identify not only topic development paths within the literature but also the most salient articles in each individual topic. The second contribution lies in forecasting the popularity of the 10 topics, based on each topic's temporal idiosyncrasies, which will assist both researchers and journal editors to select promising research topics by outlining an OSH research roadmap. This contributes to guiding a development strategy and research agenda for the post-COVID world. In the following paragraphs, we describe our sample data followed by describing the experimental results and analysis, including topic extraction, topic description, topic landscape, and topic dynamics. Finally, we conclude and discuss the implications of this research.

2 Description of the sample

We extracted article abstracts from the core collection of the Web of Science (WoS) database using the following criteria: articles published in English, whose topic terms (i.e., titles, abstracts, and keywords) include "Construction" and "Safety" or "Health" in Science Citation Index Expanded™ (SCIE) and Social Sciences Citation Index™ (SSCI) indexed journals over the period of 1991 to 2020. A total of 895 abstracts were collected, of which 662 were valid samples. Some were invalid due to the absence of an abstract or a low level of relevance to construction OSH.

Figure 1 describes the yearly distribution of articles in terms of annual article counts and the percentage of our sample article counts to the total number of SCIE and Science Citation Index (SCI) indexed articles per year (hereafter, publication percentage). The results reveal that the average annual growth rate of the publication percentage in this field reached 26.5%. Specifically, only two articles (0.003‰) were published in 1991, whereas this figure rose to 78 (0.034‰) in 2019. This demonstrates that the research field has grown significantly over the past three decades. However, the number of publications remains trivial compared to the total number of indexed articles. Our investigation also identified a systematic change in both series of article counts and the publication percentage over time. Specifically, there was a rapid development in this field between 1994 and 2001: the growth rate jumped from 0% in 1993 to 200% in 1994 and reached a peak level of 200% again in 2001.

The authors of the extracted articles are from 71 countries, especial ly the United States (US) (40%), China (11.06%), Netherlands (7.49%), Australia (6.59%), and Sweden (6.37%). Table 1 indicates that the 10 most active research institutions in our sample are Hong Kong Polytech University (3.35%), University of Michigan (3.13%), University of Washington (2.68%), National Institute for Occupational Safety and Health (2.68%), Umea University (2.46%), Karolinska Institute (2.35%), Duke University (2.23%), City University of Hong Kong (2.12%), Vrije University Amsterdam (2.12%), and Harvard University (2.12%).

Our sample articles span in 89 WoS research areas, including "Public, Environmental and Occupational Health", "Engineering", "Environmental Sciences and

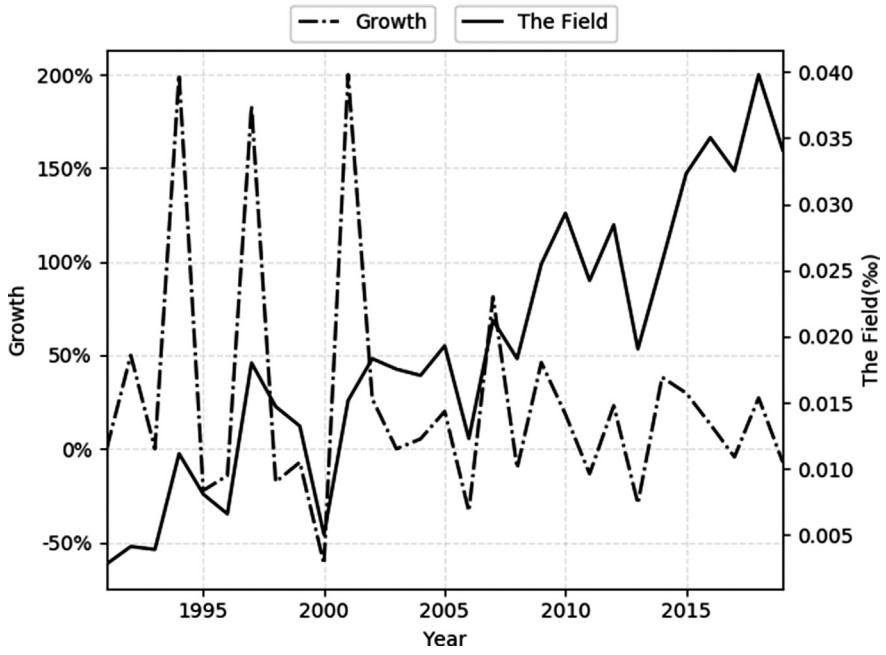

Fig. 1: The publication percentage and its growth rate of "construction health and safety".

Tab. 1: Top 10 most active research institutions

Institution	Country/region	Percentage (%)
Hong Kong Polytech University	Hong Kong, China	3.35
University of Michigan	US	3.13
University of Washington	US	2.68
National Institute for Occupational Safety and Health	US	2.68
Umea University	Sweden	2.46
Karolinska Institute	Sweden	2.35
Duke University	US	2.23
City University of Hong Kong	Hong Kong, China	2.12
Vrije University Amsterdam	Netherlands	2.12
Harvard University	US	2.12

OSH, occupational safety and health; US, United States.

Tab. 2: Top 10 research areas

Research area	#Article	Percentage (%)
Public, environmental and occupational health	493	55.08
Engineering	186	20.78
Environmental sciences and ecology	86	9.61
Construction and building technology	69	7.71
Business and economics	45	5.03
Psychology	35	3.91
Toxicology	30	3.35
General and internal medicine	28	3.13
Operations research and management science	26	2.91
Transportation	25	2.79

Ecology," and many others. Table 2 summarizes the top 10 research areas, covering 83.02% of the sample articles. These articles were published in 318 journals. Table 3 lists the 10 most popular ones.

3 Methodology and results

3.1 Topic extraction

We followed the approach of Guo et al. (2018) to classify the articles into topics using the hierarchical clustering method. First, we built a corpus containing the titles, keywords, and abstracts of all sample articles. All texts were converted to lower case and stop-words as well as punctuation based on the standard NLTK list were removed. The remaining words were converted to their stems and then a term-frequency-inverse-document-frequency (tf-idf) matrix was generated using the widely used method defined by Blei and Lafferty (2009), Grun and Hornik (2011), and Antons et al. (2016). The matrix included only the terms whose tf-idf values were just above the median of all tf-idf values of the entire vocabulary. These preprocessing procedures resulted in a Document-Term matrix for further analyses, which served as

Tab. 3: Top 10 research journals

Journal name	#Article	Percentage (%)
American Journal of Industrial Medicine	91	10.17
Journal of Occupational and Environmental Medicine	36	4.02
Journal of Construction Engineering and Management	34	3.80
Occupational and Environmental Medicine	32	3.58
International Journal of Environmental Research and Public Health	28	3.13
Scandinavian Journal of Work Environment and Health	26	2.91
Safety Science	25	2.79
International Archives of Occupational and Environmental Health	19	2.12
Automation in Construction	17	1.90
Journal of Safety Research	16	1.79

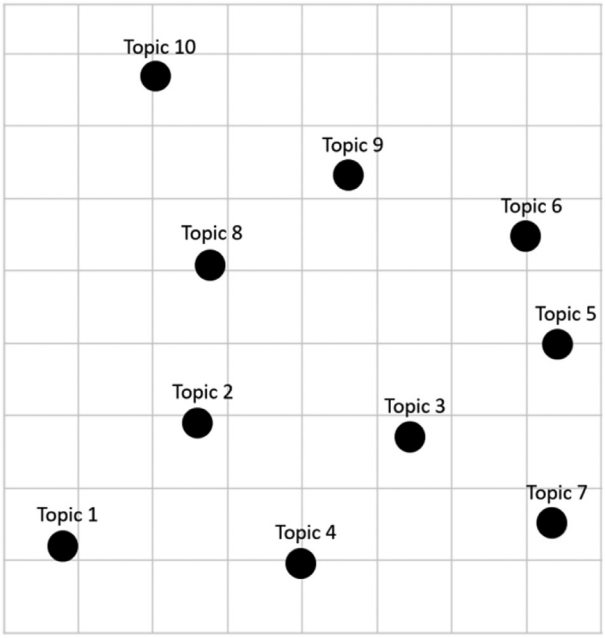

Fig. 2: Inter-topic distance in a two-dimensional space using MDS. MDS, multidimensional scaling.

the input of the Ward clustering algorithm and a cosine-distance-based hierarchical dendrogram was generated. The articles were roughly classified into 10 topic clusters.

To make sure about the appropriate number of topics, we followed the conceptual content analysis approach outlined by Hsieh and Shannon (2005), to identify the most outstanding topics in the research field of construction health and safety. We randomly split the 662 valid sample article abstracts into three sets: 220 abstracts in the first set as the "training samples," 220 abstracts in the second set as the "validation sample," and the remaining 222 abstracts in the third set as the "test sample." We invited 20 student volunteers to read the full text of the first set of article abstracts. The students and the author team worked together to develop sets of terms and coding rules. The team of students and authors developed a set of 10 topics from the first set based on these terms. The team found that it was appropriate to classify the second set of 220 abstracts into the same 10 topics. These findings indicated that 10 could be a good number of topics, which allows us to categorize all sample article abstracts optimally. To verify the robustness of the optimal number of topics, the team tried to classify the third set of articles into the same 10 topics. The team found that these 10 topics well accommodated the articles in the third set. Therefore, we concluded that there were 10 topics emerged from the sampled articles.

Finally, each student and author read the abstract of the 662 articles and assigned a topic loading (i.e., a score between 0 and 1 that refers to the likelihood that a particular article belongs to a particular topic) to every topic. The topic loading of a topic is assigned as 0 if an article abstract is irrelevant to a particular topic. In this way, we generated an article-topic matric with a size of 662*10, in which each row represents an article and each column represents a topic. Each article was assigned to the dominant topic with the highest topic loading. To assess the topic extraction quality, we used the multidimensional scaling (MDS) method to plot the inter-topic distances in a two-dimensional space. As shown in Figure 2, the 10 topics do not cluster but spread almost evenly.

3.2 Topic description

In the previous section, we extracted 10 distinct topics from the sample articles. In order to draw an in-depth understanding of the current literature of construction OSH, we generated a label for each of these topics followed by distilling opinions based on prior research works (Table 4).

3.3 Topic landscape

We followed Guo et al. (2018) to use the Herfindahl–Hirschman Index (HHI) to explore the topic diversity. Specifically, HHI is widely used by economists and

Tab. 4: Description of the extracted 10 distinct topics

Topic ID	Topic label	Key features	Description
1	Physical health and disease	Health status and various related indicators. Correlation of strain-related diseases and other diseases.	Construction workers need more education on nutrition, healthy behaviors, and workplace injury prevention. Training and education programs can promote nutrition and health and safety knowledge such that the dietary behavior of construction workers can be improved, and workplace injuries can be reduced. Existing studies demonstrate that obesity and high physical workload have a synergistic, negative effect on work ability.
2	Mental health and dysphrenia	Mental health problems, work pressure, suicidal behavior, and mental disorders.	Management should focus on improving construction workers' engagement with the workplace and work/life balance. Many male construction workers do not seek out help despite experiencing mental health issues. Active interventions and self-compassion are important and can medicate and adjust workers' adverse emotional and mental states.
3	Gender and age	Gender ratio Differences in personal income due to gender and age. Aging of the labor market Labor replacement and updating.	Trade-specific age trends within the construction industry reveals that age is a boundary condition of job resources stress link. In addition, female construction workers are exposed to various safety hazards as they hardly find properly fitting Personal Protective Equipment (PPE). Understanding the needs and motivations of construction employees of different genders and ages is crucial to their health and safety and well-being.
4	Migrant and race	Income difference between ethnicities and immigrants. Behavior and awareness of health and safety.	The construction industry is multicultural, containing a large number of migrants. Effective communication on construction sites is crucial to improve construction performance. Cultural barriers affect the clarity of communication between construction workers, which in turn, affect the jobsite safety and productivity. Understanding and managing various cultures and improving work practice can help improve the health and safety, and the economic and social construction of migrant workers.
5	Vocational ability and training	Degree of skill proficiency. Improvement and importance of vocational skill training for health and safety.	Jobsite accident rate remains high despite substantial efforts have spent in improving the health and safety education, Existing one-size-fits-all training program does not cater to the needs to construction workers who are heterogeneous in nature. The major factor contributing to ineffective training is the information delivery gaps during the knowledge-transfer process, i.e., ineffectiveness of transferring their knowledge and skills to job sites. Personalized training with active interventions is a promising approach for a diverse group of construction workers.
6	Occupational exposure and accident	Effects of occupational exposure on health and safety. Workplace accidents and work-related injuries.	The construction industry is a dangerous workplace. Construction workers are highly likely to exposed to reactive chemicals, solar Ultraviolet (UV) radiation, poor safety environment and PPE, stress, and other health and safety hazards, which cause occupational injuries. However, most of the factors significantly associated with these dangers and injuries can be prevented and controlled. It is essential to increase workers' health and safety awareness and implement comprehensive exposure control programs, including regulations in regional and national legislation, in order to minimize accidents and injuries.
7	Safety climate and management	Knowledge of the construction safety. Relationship between management and workers Management strategies and methodologies.	Different factors impact the health and safety of jobsite construction workers. Prior studies explored whether construction workers' social capital affects their safety behaviors and showed that structural and relational dimensions of workers' social capital have a significant and direct effect on workers' safety behaviors, whereas the cognitive dimensions does not. Effective and trustful interpersonal relationships between construction workers and management can improve workers' safety behaviors. Organizational factors have the greatest impact on construction workers' safety while human factors have the least.

(*Continued*)

Tab. 4: Continued

Topic ID	Topic label	Key features	Description
8	Job situation, security and policies	Working conditions, job security, pensions, subsidies and other welfare benefits provided by management.	Jobsite facilities and labor welfare benefits directly or indirectly affect the productivity of the construction workforce.
9	Tobacco, alcohol and drugs	Addiction to smoking, alcohol, and drugs. Quitting behavior.	The construction is a generally male-dominated industry that is more susceptible to tobacco, alcohol, and drug use, which pose a serious risk to construction workers' health and safety. The influence of tobacco, alcohol, and drug workplace availability and norms highlight the importance of workplace culture and health. Hybrid and multifaceted approaches involving culture change and personal-level responses targeting vulnerable, low-acculturation-level, and migrant construction workers are necessary in order to minimize risk from smoking-, alcohol-, and drug-related harm and effects. Specific tobacco, alcohol, and drug treatments tailored for vulnerable and migrant workers' culture are essential to support sustainable worksite cessation efforts and can be integrated into other health and safety programs.
10	Smart devices	Use of smart products. Use of intelligent monitoring and management devices. Intelligent management systems.	One major reason for unsatisfactory occupational health and safety records in the construction industry is attributed to high physical requirements and dynamic and hostile working environments. Construction workers are subjected to ergonomic risks and other potential injuries. Traditional jobsite worker monitoring relies on construction workers' self-reporting and subjective questionnaires. Both intrusive (e.g., wearable devices and attached sensors) and non-intrusive (e.g., computer-vision-based motion capturing) methods are used to monitor workers' status by effectively and objectively collecting their physical and mental data and then used to access workers' physical states.

government regulators to measure the level of diversity. HHI refers to the sum of squares of the total revenue or the percentage of the total assets of each market competitor in an industry. HHI is commonly used to calculate the changes in the market share, namely, the dispersion degree of firm size in the market. The value of HHI varies between $\frac{1}{N}$ and 1. Higher values of the index indicate a higher market concentration and monopoly power. Analogically, in the context of topic modeling, Guo et al. (2018) define HHI as the sum of squares of the ratio of the score of each article to the total score. The value of HHI for each topic still varies between $\frac{1}{N_i}$ and 1, where N_i is the number of articles with a nonzero score in topic i. Note that the sample size of N_i is different for each topic. We normalized HHI for each topic in order to make a fair horizontal comparison of concentration distribution between different topics. Specifically,

$$H_i = \frac{N_i \sum_{i=1}^{N_i} \left(\frac{Y_{ij}}{Y_i}\right)^2 - 1}{N_i - 1} * 100\%,\ i = 1,2,\ldots,10 \quad (1)$$

where

Y_{ij}: the score of an article j in the topic i,
Y_i: the total score of topic i,
N_i: number of articles with a nonzero score of topic i.

Table 5 illustrates the score of the normalized HHI (i.e., H_i) for all the 10 topics. The number of articles per topic ranges from 14 to 193 (mean = 66.2 and a standard deviation = 65.3). The three most prevalent topics are "Physical Health and Disease" (Topic 1), "Occupational Exposure and Accident" (Topic 6), and "Safety Climate and Management" (Topic 7), while the three least prevalent topics are "Gender and Age" (Topic 3), "Migrant and Race" (Topic 4), and "Tobacco, Alcohol and Drugs" (Topic 9). Table 5 also provides the number of articles of each dominant topic and the H_i results (in percentage) in the range length of $\left(\left[\left[\frac{1}{N},1\right]\right]\right)$ with their standard deviations σ. An article contains diverse topics if its H_i value is small, whereas it contains a salient topic if its H_i value is large. As shown, some topics exhibited a high degree of concentration. For example, "Gender and Age" (Topic 3, $H_3 = 0.65\%$), "Migrant and Race" (Topic 4, $H_4 = 0.80\%$), and "Tobacco, Alcohol and Drugs" (Topic 9, $H_9 = 0.51\%$) are the three topics with the highest H_i values. They are all topics with

Tab. 5: Topic Label and H_i values for the 10 topics

Topic ID	Topic label	# Article (%)	H_i (%)	σ
1	Physical health and disease	193 (29.15)	0.06	0.40
2	Mental health and dysphrenia	41 (6.19)	0.34	0.22
3	Gender and age	15 (2.27)	0.65	0.11
4	Migrant and race	14 (2.11)	0.80	0.11
5	Vocational ability and training	29 (4.38)	0.42	0.18
6	Occupational exposure and accident	183 (27.64)	0.05	0.42
7	Safety climate and management	102 (15.41)	0.10	0.33
8	Job situation, security and policies	28 (4.23)	0.45	0.18
9	Tobacco, alcohol and drugs	26 (3.93)	0.51	0.17
10	Smart devices	31 (4.68)	0.21	0.20

Tab. 6: The most popular authors per topic

Topic ID	Authors (Surname, First name)
1	van der Beek, Allard J.; Brenner, H; Proper, Karin I.
2	Gullestrup, Jorgen; LaMontagne, Anthony D.; Milner, Allison
3	Chan, Alan H. S.; Dong, Xiuwen Sue; Wang, Xuanwen
4	Evia, Carlos; Dietz, Noella A.; Asfar, Taghrid
5	Hess, Jennifer A.; Weeks, Douglas L.; Anton, Dan
6	Dong, Xiuwen Sue; Jarvholm, Bengt; Heederik, Dick
7	Chan, Albert P. C.; McCabe, Brenda; Li, Shuquan
8	Jarvholm, Bengt; Stattin, Mikael; Burdorf, Alex
9	Ye, Weimin; Duraisingam, Vinita; Pidd, Ken
10	Lee, SangHyun; Li, Heng; Jebelli, Houtan

strong topicality, resulting in a relatively concentrated distribution of scores. In contrast, topics exhibiting the lowest H_i values include "Physical Health and Disease" (Topic 1, $H_1 = 0.06\%$), "Occupational Exposure and Accident" (Topic 6, $H_6 = 0.05\%$), and "Safety Climate and Management" (Topic 7, $H_7 = 0.10\%$). These articles contain a wide range of subtopics with sparse distributions, rending their topic σ large.

We indicated the top authors and journals that appear the most frequently on each topic in Tables 6 and 7, respectively, to identify the communities that have contributed the most to each topic in the field of construction OSH. The results show that Sue Xiuwen Dong from the center for construction research and training in the United States is one of the most active and influential authors in Topics 3 and 6. Likewise, Jarvholm Bengt is among the most productive and important authors in Topics 6 and 8. The results also show that the *American Journal of Industrial Medicine* is the most popular journal for eight topics. Table 8 summarizes the three most representative and recent articles (i.e., with the highest topic loading) of each topic published in the top journals.

Tab. 7: The most popular journals per topic

Topic ID	Journals
1	*American Journal of Industrial Medicine* *Journal of Occupational and Environmental Medicine* *Scandinavian Journal of Work Environment and Health*
2	*International Journal of Environmental Research and Public Health* *Journal of Construction Engineering and Management* *Engineering Construction and Architectural Management*
3	*American Journal of Industrial Medicine* *International Journal of Environmental Research and Public Health* *Journal of Occupational and Environmental Medicine*
4	*American Journal of Industrial Medicine* *Journal of Construction Engineering and Management* *Journal of Occupational and Environmental Medicine*
5	*American Journal of Industrial Medicine* *Safety Science* *Journal of Occupational and Environmental Medicine*
6	*American Journal of Industrial Medicine* *Occupational and Environmental Medicine* *Journal of Occupational and Environmental Medicine*
7	*Journal of Construction Engineering and Management* *Safety Science* *American Journal of Industrial Medicine*
8	*American Journal of Industrial Medicine* *Journal of Occupational and Environmental Medicine* *Journal of Construction Engineering and Management*
9	*American Journal of Industrial Medicine* *International Journal of Cancer* *Journal of Occupational and Environment Medicine*
10	*Automation in Construction* *Journal of Construction Engineering and Management* *Journal of Computing in Civil Engineering*

3.4 Topic dynamics

As discussed earlier (Figure 1), the number of articles published in the field of construction OSH has increased steadily since 1991. We generated a bubble chart in Figure 3 to illustrate the development of each topic over these years more intuitively. The size of the bubbles is proportional to the number of published articles. Some topics experienced rapid growth. Specifically, Topic 1

"Physical Health and Disease" increases from 0 papers in 1990 to 22 papers in 2017. Topic 6 "Occupational Exposure and Accident" increases from 2 papers in 1990 to 22 in 2019. Topic 7 "Safety Climate and Management" increases from 0 papers in 1990 to 16 in 2018. Some topics, for example, Topic 10 "Smart Devices", exhibit an escalation in recent years and future potential popularity, although the number of published articles is not significant.

Tab. 8: Topic 3 representative articles per topic

Topic ID	Articles
1	Tonnon et al. (2019); Chung et al. (2019); Caban-Martinez et al. (2018)
2	Campbell and Gunning (2020); Kotera et al. (2019); Xing et al. (2019)
3	Onyebeke et al. (2016); Sokas et al. (2019); Yaldiz et al. (2018)
4	Al-Bayati et al. (2017); Dong et al. (2010); Khan and Sandhu (2016)
5	Hussain et al. (2020); Khokhar et al. (2019); Xu et al. (2019)
6	Bello et al. (2020); Kiconco et al. (2019); Moldovan et al. (2020)
7	Alomari et al. (2020); Li et al. (2020); Marín et al. (2019)
8	Gupta et al. (2018); Roelen et al. (2014); Schofield et al. (2019)
9	Asfar et al. (2019); Chapman et al. (2020); Roche et al. (2020)
10	Osunsanmi et al. (2019); Yu et al. (2019a); Yu et al. (2019b)

Next, we aim to perform time series forecasting analysis to predict the future research direction in the next decade and make a comparison between topics. We followed Hyndman and Athanasopoulos (2018) to use Autoregressive Integrated Moving Average (ARIMA), which focuses on depicting the autocorrelation and is known to be a simple yet powerful method to time series forecasts (Geurts et al. 1977). The "AR" part is defined as a linear regression on the previous time series values, while the "MA" part is deemed as a linear regression of the current values of the series against prior random shocks. The "I" (i.e., Integrated) part means that the data values are replaced with the difference between their values and one or several previous values, allowing nonstationary series to be modeled (Geurts et al. 1977).

Following the commonly used Box-Jenkins methodology (Geurts et al. 1977), we conducted 11 time series forecasting analysis (i.e., the field integrally and 10 topics, respectively). Timing data predicted by ARIMA should be stable without trends or periodicity. This means that the mean value of the timing data has a constant amplitude on the time axis, and its variance tends to be the same stable value on the time axis. We used the Augmented Dickey-Fuller (ADF) (Said and Dickey, 1984) test to investigate the stability of time series data. ADF is a widely used unit root test method. A difference is

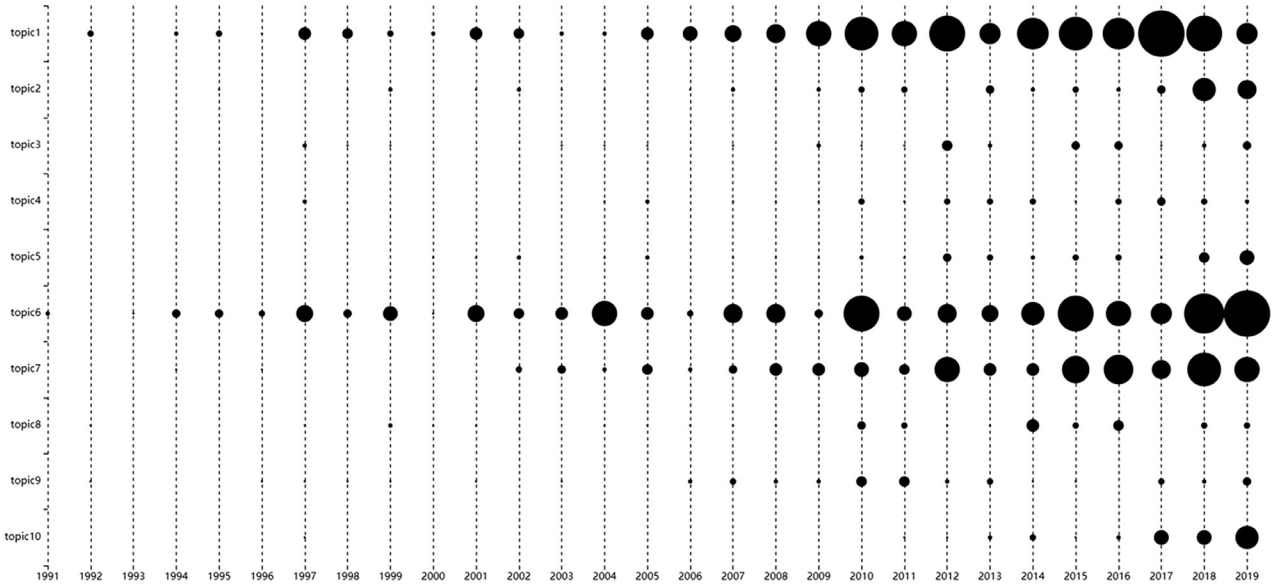

Fig. 3: Topics development over the past three decades.

Tab. 9: ARIMA results and forecasting

Topic	Order[1]	logLik	AIC	BIC	HQIC[2]	RMSE	LBstat[3]	Avg Future Gth (%)	Category
The field	(0,1,1)	−93.83	193.65	197.76	194.94	6.14	0.53	2.57	Benchmark
1	(0,1,1)	−69.63	145.26	149.14	146.41	3.16	0.88	8.65	Hot
2	(0,1,1)	−42.77	91.54	94.67	92.22	1.85	0.99	3.62	Hot
3	(0,1,1)	−27.98	61.95	64.45	62.20	1.22	0.87	0.91	Stable
4	(0,1,1)	−21.71	49.43	51.74	49.54	0.92	0.57	7.85	Hot
5	(0,2,2)	−31.74	71.47	75.03	71.96	1.26	0.85	7.56	Hot
6	(0,2,4)	−78.14	164.28	169.61	165.91	3.66	0.21	3.52	Hot
7	(0,1,2)	−45.62	97.25	100.23	97.83	2.29	0.09	4.95	Hot
8	(0,0,0)	−34.89	73.78	75.67	74.10	1.52	0.14	−2.98	Cold
9	(0,2,1)	−34.48	76.95	80.94	77.73	1.27	0.92	1.70	Stable
10	(0,2,1)	−19.84	47.67	48.88	46.34	1.57	0.81	15.74	Hot

[1]The order is listed as p, d, q.
[2]HQIC: Hannan-Quinn information criterion.
[3]LBstat: Ljung-Box portmanteau statistic.
ARIMA, autoregressive integrated moving average.

formulated to eliminate the periodic factors for the time series data that do not meet the requirements of stationarity. The parameter d presents the degree of differences to make the data stable. The parameters p and q present the autocorrelation function (ACF) and partial autocorrelation function (PACF), respectively (Hyndman and Athanasopoulos, 2018). We conducted a grid-search to compute the scores of Root Mean Square Error (RMSE), Akaike information criterion (AIC), and Bayesian information criterion (BIC) with different combinations of orders (Hyndman and Athanasopoulos, 2018). The order specification with the lowest scores was returned. To verify the quality of the model, we plotted its residual to see whether it appears as entirely random white noise and conducted the Ljung-Box test to formally check whether the errors were uncorrelated across many lags (Hyndman and Athanasopoulos, 2018). Otherwise, we improved the model by removing all the remaining trends.

Optimized ARIMA models were used to forecast the publication percentage of the first and of each topic for the next 5 years (i.e., 2021–2025). We use the predicted average annual growth (i.e., Avg Future Gth) rate as the indicator of future topic prevalence (Table 9). In general, the field may expand steadily but its annual growth rate would become slow (2.57%). This Avg Future Gth is far lower compared to the peak annual growth rate of 200% in 1994 and 2001. This indicates that the construction OSH will stay as an active research field, whereas the traditional topics are unlikely to attract explosive attention in the future. The 10 topics were categorized into three types of trends, i.e., hot, stable, and cold. Specifically, hot topics are topics whose forecast annual growth rates are higher than or equal to that of the field. Stable topics are topics whose rates are positive or equal to 0 but smaller than that in the field. Cold topics are topics whose rates are negative. Table 9 shows that Topic 8 is a cold topic, Topics 3 and 9 are stable topics, while the rest remain as hot topics.

4 Discussion and conclusions

In order to provide a systematic review of the state of research of the construction OSH over the past three decades, this study maps the topic landscape while identifying the salient development trajectories. We first introduced our samples by describing the publication development over this period. Then, we provided the most popular research organizations, research areas, and journals in construction OSH. Next, we implemented a content analysis approach to classify our samples into 10 distinct topics followed by measuring the inter-topic distances using the MDS algorithm. To understand the hidden structure of the extracted 10 different topics covering the overall development of this field, we used the HHI index to measure the topic diversity level. A dynamic analysis was performed revealing the complex relationship of the content structure in the construction OSH. The forecasting results suggested that although the construction OSH might remain as an active research direction, some topics might lose their popularity (e.g., Topic 8, Job Situation, Security, and Policies) or stay flat (e.g., Topic 3, Gender and Age and Topic 9, Tobacco, Alcohol, and Drugs) in the future.

In particular, the proposed ARIMA model performed time series forecasting analysis to predict that

"Job Situation, Security, and Policies" (Topic 8, –2.98%) might shrink in the next 5 years. On one hand, it is not surprising to see that this topic shows a downtrend as a policy of job situation and security and does not change as fast as technology. However, this is uncertain and highly likely to go uptrend in the upcoming years due to the current global crisis of COVID-19, which has drastically changed the world, including the construction industry. It is almost certain that the construction industry will roll out new policies and codes of practice of construction OSH for practitioners in the post-COVID period. Likewise, COVID-19 has raised people's health awareness such that the use of tobacco, alcohol, and drugs might be reduced in the future. This is also due to the rapid development of digital technology, which promotes the development of related policies and regulations while putting forward higher standards for cyber security. Additionally, gender and age gaps would be further bridged in the digital era.

The other seven topics were predicted to have positive growth rates. Specifically, "Physical Health and Disease" (Topic 1, 8.65%), "Migrant and Race" (Topic 4, 7.85%), "Vocational Ability and Training" (Topic 5, 7.56%), "Safety Climate and Management" (Topic 7, 4.95%), and "Smart Devices" (Topic 10, 15.74%) will maintain popularity. Of these topics, Topic 10 "Smart Devices" is considered to be the most promising topic in the future. With the rapid development of big data, the Internet of Things, and high-performance computing capabilities, digital technology has been greatly implemented in various scenarios in the construction industry. Most importantly, it has been and will be an internationally important area – arguably even more relevant in the post-COVID world. This includes but is not limited to the digital twin technology, Building Information Modelling (BIM), Virtual Reality (VR)/Augmented Reality (AR), remote sensing technology, on-site tracking/inspection, wearable devices, mobile reporting, e-learning/training, and drone. Accelerating the digital transformation will be the best way to improve the jobsite environment and management, develop a culture where health, safety, and wellbeing is at its heart for construction workers, and mitigate the human and capital costs of accidents/occupational diseases at work. It is worth noting that Topic 1 "Physical Health and Disease" will maintain its popularity in the future. This is not surprising as the construction industry remains to be one of the most dangerous and stressful workplaces. Likewise, Topic 4, "Migrant and Race," keeps growing positively because the construction industry keeps absorbing lots of migrants with different cultures and backgrounds. Topic 5 "Vocational Ability and Training" and Topic 7 "Safety Climate and Management" also keep growing because skill training, and certain social and management factors affect workers' safety behaviors and their health.

This study investigates the most relevant WoS literature, providing researchers and practitioners a systematic overview of the field of the construction OSH. Compared to other similar studies, the contribution of this study is twofold. Not only does it uncover the hidden structure of distinct topics and development trajectories in related abstracts of a body of most representative OSH articles, but it also enables readers to examine the detailed profile of each topic, estimates its relative salience, and suggests the future trend. It assists construction firms to identify and invest new technologies of construction OSH. By describing the body of knowledge at a relatively granular level, this study contributes to a rich understanding of the topic landscape of this field. Similarly, it also helps journal editors to compare the field's current topic landscape against their journal's editorial board or promoted through special issues.

Although this study provides valuable insights, there are some limitations. The samples used in this study were collected from WoS. To our best knowledge, it is the single most authoritative source for "high-impact" publications and has relatively better coverage of the research field of this study than other academic databases. However, WoS covers mainly mainstream journals and articles, especially those in English. As a result, the analysis provided in this study does not take into account articles published in emerging journals, in non-English language ones, or other types of publications (e.g., books, conference papers, technical reports, or dissertations). In addition, this study does not specify forecasting models with any external bibliometric factors that may correlate with the growth or decline of a topic time series. Therefore, for future research, we plan to collect publication records from other open-access online publications. We also plan to investigate bibliometric determinants of topic dynamics.

Acknowledgements

The study is supported by the Key Research and Development Project of Shandong Province (2018GGX101007). The funding body had no role in the study design, data collection, analyses, and interpretation, or in writing the manuscript.

References

Al-Bayati, A. J., Abudayyeh, O., Fredericks, T., & Butt, S. E. (2017). Managing cultural diversity at U.S. construction sites: Hispanic workers' perspectives. *Journal of Construction Engineering and Management*, *143*(9), pp. 04017064. doi: 10.1061/(ASCE)CO.1943-7862.0001359.

Alomari, K., Gambatese, J., Nnaji, C., & Tymvios, N. (2020). Impact of risk factors on construction worker safety: A Delphi Rating Study based on field worker perspective. *Arabian Journal for Science and Engineering*, *45*(10), pp. 8041-8051. doi: 10.1007/s13369-020-04591-7.

Antons, D., Kleer, R., & Salge, T. O. (2016). Mapping the topic landscape of JPIM, 1984–2013: In search of hidden structures and development trajectories. *Journal of Product Innovation Management*, *33*(6), pp. 726-749. doi: 10.1111/jpim.12300.

Antwi-Afari, M. F., Li, H., Edwards, D. J., Pärn, E. A., Seo, J., & Wong, A. Y. (2017). Biomechanical analysis of risk factors for work-related musculoskeletal disorders during repetitive lifting task in construction workers. *Automation in Construction*, *83*, pp. 41-47. doi: 10.1016/j.autcon.2017.07.007.

Arcury, T. A., Grzywacz, J. G., Chen, H., Mora, D. C., & Quandt, S. A. (2014). Work organization and health among immigrant women: Latina manual workers in North Carolina. *American Journal of Public Health*, *104*(12), pp. 2445-2452. doi: 10.2105/AJPH.2013.301587.

Arcury, T. A., Mills, T., Marín, A. J., Summers, P., Quandt, S. A., Rushing, J., et al. (2012). Work safety climate and safety practices among immigrant Latino residential construction workers. *American Journal of Industrial Medicine*, *55*(8), pp. 736-745. doi: 10.1002/ajim.22058.

Asfar, T., McClure, L. A., Arheart, K. L., Ruano-Herreria, E. C., Gilford, C. G., Moore, K., et al. (2019). Integrating worksite smoking cessation services into the construction sector: Opportunities and challenges. *Health Education and Behavior*, *46*(6). pp. 1024-1034. doi: 10.1177/1090198119866900.

Bello, A., Xue, Y., Gore, R., Woskie, S., & Bello, D. (2020). Exposures and urinary biomonitoring of aliphatic isocyanates in construction metal structure coating. *International Journal of Hygiene and Environmental Health*, *226*, pp. 113495. doi: 10.1016/j.ijheh.2020.113495.

Blei, D. M., & Lafferty, J. D. (2009). Topic models, text mining: Classification, clustering, and applications. pp. 71–89. doi: 10.1145/1143844.1143859.

Caban-Martinez, A. J., Moore, K. J., Clarke, T. C., Davila, E. P., Clark, J. D., Lee, D. J., et al. (2018). Health promotion at the construction work site: The Lunch Truck Pilot Study. *Workplace Health and Safety*, *66*(12), pp. 571-576. doi: 10.1177/2165079918764189.

Campbell, M. A., & Gunning, J. G. (2020). Strategies to improve mental health and well-being within the UK construction industry. *Proceedings of Institution of Civil Engineers: Management, Procurement and Law*, *173*(2), pp. 64-74. doi: 10.1680/jmapl.19.00020.

Chapman, J., Roche, A. M., Duraisingam, V., Phillips, B., Finnane, J., Pidd, K. (2020). Working at heights: Patterns and predictors of illicit drug use in construction workers. *Drugs: Education, Prevention and Policy*, *28*(1), pp. 67-75. doi: 10.1080/09687637.2020.1743645.

Cheng, T., Migliaccio, G. C., Teizer, J., & Gatti, U. C. (2013). Data fusion of real-time location sensing and physiological status monitoring for ergonomics analysis of construction workers. *Journal of Computing in Civil Engineering*, *27*(3), pp. 320-335. doi: 10.1061/(ASCE)CP.1943-5487.0000222.

Chung, L. M., Chung, J. W., & Chan, A. P. (2019). Building healthy eating knowledge and behavior: An evaluation of nutrition education in a skill training course for construction apprentices. *International Journal of Environmental Research and Public Health*, *16*(23), pp. 4852. doi: 10.3390/ijerph16234852.

Dong, X. S., Men, Y., & Ringen, K. (2010). Work-related injuries among Hispanic construction workers-evidence from the Medical Expenditure Panel Survey. *American Journal of Industrial Medicine*, *53*(6), pp. 561-569. doi: 10.1002/ajim.20799.

Finneran, A., & Gibb, A. G. (2013). W099: Safety and health in construction: Research Roadmap - Report for Consultation. *CIB Publication*, *376*. doi: 10.1016/j.tetlet.2005.02.138.

Geurts, M., Box, G. E. P., & Jenkins, G. M. (1977). Time series analysis: Forecasting and control. *Journal of Marketing Research*. *14*(000002), pp. 269. doi: 10.2307/3150485.

Grün, B., & Hornik, K. (2011). Topicmodels: An R package for fitting topic models. *Journal of Statistical Software*, *40*(13), pp. 30571. doi: 10.18637/jss.v040.i13.

Guo, L., Li, S., Lu, R., Yin, L., Gorson-Deruel, A., & King, L. (2018). The research topic landscape in the literature of social class and inequality. *PLoS One*, *13*(7), pp. e0199510. doi: 10.1371/journal.pone.0199510.

Gupta, M., Hasan, A., Jain, A. K., & Jha, K. N. (2018). Site amenities and workers' welfare factors affecting workforce productivity in Indian Construction Projects. *Journal of Construction Engineering and Management*, *144*(11), pp. 04018101. doi: 10.1061/(ASCE)CO.1943-7862.0001566.

HSE. (2019). Construction statistics in Great Britain. *Annual Statistics,* Available at https://www.hse.gov.uk/sTATIsTICs/industry/construction.pdf [accessed 30 July, 2020].

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), pp. 1277-1288. doi: 10.1177/1049732305276687.

Hussain, R., Pedro, A., Lee, D. Y., Pham, H. C., & Park, C. S. (2020). Impact of safety training and interventions on training-transfer: Targeting migrant construction workers. *International Journal of Occupational Safety and Ergonomics*, *26*(2), pp. 272-284. doi: 10.1080/10803548.2018.1465671.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice*. 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2. [accessed 30 July, 2020].

ILO. (2001). Recruitment practices pose problems for construction industry. Available at https://www.ilo.org/global/about-the-ilo/newsroom/news/WCMS_007879/lang – en/index.htm. [accessed 30 July, 2020].

ILO. (2003). Safety in numbers. International Labour Office. Available at https://www.ilo.org/legacy/english/protection/safework/worldday/report_eng.pdf. [accessed 3 August, 2021].

Jebelli, H., Choi, B., & Lee, S. H. (2019). Application of wearable biosensors to construction sites. I: Assessing workers' stress. *Journal of Construction Engineering and Management*, *145*(12), pp. 04019079. doi: 10.1061/(ASCE)CO.1943-7862.0001729.

Khan, A., & Sandhu, M. (2016). Benchmarking national culture and decent work practice indicators in project-based industry: Lessons from United Arab Emirates. *Benchmarking: An International Journal*, *23*(3), pp. 490-518. doi: 10.1108/BIJ-02-2014-0015.

Khokhar, M., Hou, Y., Sethar, I., Amin, W., & Shakib, M. (2019). Occupational health & safety implementation framework for pakistani construction industry in Sindh province. *3C Tecnología_Glosas de innovación aplicadas a la pyme*, pp. 253-285. doi: 10.17993/3ctecno.2019.specialissue3.253-285.

Kiconco, A., Ruhinda, N., Halage, A. A., Watya, S., Bazeyo, W., Ssempebwa, J. C., et al. (2019). Determinants of occupational injuries among building construction workers in Kampala City, Uganda. *BMC Public Health*, *19*(1), pp. 1-1. doi: 10.1186/s12889-019-7799-5.

Kotera, Y., Green, P., & Sheffield, D. (2019). Mental health shame of UK construction workers: Relationship with masculinity, work motivation, and self-compassion. *Revista de Psicologia del Trabajo y de las Organizaciones*, *35*(2), pp. 135-143. doi: 10.5093/jwop2019a15.

Lee, W., Lin, K. Y., Seto, E., & Migliaccio, G. C. (2017). Wearable sensors for monitoring on-duty and off-duty worker physiological status and activities in construction. *Automation in Construction*, *83*, pp. 341-353. doi: 10.1016/j.autcon.2017.06.012.

Li, S., Wu, X., Wang, X., & Hu, S. (2020). Relationship between social capital, safety competency, and safety behaviors of construction workers. *Journal of Construction Engineering and Management*, *146*(6), pp. 04020059. doi: 10.1061/(ASCE)CO.1943-7862.0001838.

Marín, L. S., Lipscomb, H., Cifuentes, M., & Punnett, L. (2019). Perceptions of safety climate across construction personnel: Associations with injury rates. *Safety Science*, *118*, pp. 487-496. doi: 10.1016/j.ssci.2019.05.056.

Moldovan, H. R., Wittlich, M., John, S. M., Brans, R., Tiplica, G. S., Salavastru, C., et al. (2020). Exposure to solar UV radiation in outdoor construction workers using personal dosimetry. *Environmental Research*, *181*, pp. 108967. doi: 10.1016/j.envres.2019.108967.

Onyebeke, L. C., Papazaharias, D. M., Freund, A., Dropkin, J., McCann, M., Sanchez, S. H., et al. (2016). Access to properly fitting personal protective equipment for female construction workers. *American Journal of Industrial Medicine*, *59*(11), pp. 1032-1040. doi: 10.1002/ajim.22624.

Osunsanmi, T. O., Oke, A. E., & Aigbavboa, C. O. (2019). Survey dataset on fusing RFID with mobile technology for efficient safety of construction professionals. *Data in Brief*, *25*, pp. 104290. doi: 10.1016/j.dib.2019.104290.

Roche, A. M., Chapman, J., Duraisingam, V., Phillips, B., Finnane, J., Pidd, K. (2020). Construction workers' alcohol use, knowledge, perceptions of risk and workplace norms. *Drug and Alcohol Review*, *39*(7), pp. 941-949. doi: 10.1111/dar.13075.

Roelen, C. A., Heymans, M. W., Twisk, J. W., van der Klink, J. J., Groothoff, J. W., van Rhenen, W. (2014). Work ability index as tool to identify workers at risk of premature work exit. *Journal of Occupational Rehabilitation*, *24*(4), pp. 747-754. doi: 10.1007/s10926-014-9505-x.

Said, S. E., & Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, *71*, pp. 599-607. doi: 10.1093/biomet/71.3.599.

Schofield, K., Ryan, A. D., & Dauner, K. N. (2019). Comparing disability and return to work outcomes between alternative and traditional workers' compensation programs. *American Journal of Industrial Medicine*, *62*(9), pp. 755-765. doi: 10.1002/ajim.23017.

Siu, O. L., Phillips, D. R., & Leung, T. W. (2004). Safety climate and safety performance among construction workers in Hong Kong: The role of psychological strains as mediators. *Accident Analysis and Prevention*, *36*(3), pp. 359-366. doi: 10.1016/S0001-4575(03)00016-2.

Sokas, R. K., Dong, X. S., & Cain, C. T. (2019). Building a sustainable construction workforce. *International Journal of Environmental Research and Public Health*, *16*(21), pp. 4202. doi: 10.3390/ijerph16214202.

Tonnon, S. C., Robroek, S. R., van der Beek, A. J., Burdorf, A., van der Ploeg, H. P., Caspers, M., et al. (2019). Physical workload and obesity have a synergistic effect on work ability among construction workers. *International Archives of Occupational and Environmental Health*, *92*(6), pp. 855-864. doi: 10.1007/s00420-019-01422-7.

Xing, X., Li, H., Li, J., Zhong, B., Luo, H., & Skitmore, M. A. (2019). A multicomponent and neurophysiological intervention for the emotional and mental states of high-altitude construction workers. *Automation in Construction*, *105*, pp. 102836. doi: 10.1016/j.autcon.2019.102836.

Xu, S., Zhang, M., & Hou, L. (2019). Formulating a learner model for evaluating construction workers' learning ability during safety training. *Safety Science*, *116,* pp. 97-107. doi: 10.1016/j.ssci.2019.03.002.

Yaldiz, L. M., Truxillo, D. M., Bodner, T., & Hammer, L. B. (2018). Do resources matter for employee stress? It depends on how old you are. *Journal of Vocational Behavior*, *107*, pp. 182-194. doi: 10.1016/j.jvb.2018.04.005.

Yi, W., Chan, A. P., Wang, X., & Wang, J. (2016). Development of an early-warning system for site work in hot and humid environments: A case study. *Automation in Construction*, *62*, pp. 101-113. doi: 10.1016/j.autcon.2015.11.003.

Yu, Y., Li, H., Yang, X., Kong, L., Luo, X., & Wong, A. Y. (2019a). An automatic and non-invasive physical fatigue assessment method for construction workers. *Automation in Construction*, *103*, pp. 1-2. doi: 10.1016/j.autcon.2019.02.020.

Yu, Y., Yang, X., Li, H., Luo, X., Guo, H., & Fang, Q. (2019b). Joint-level vision-based ergonomic assessment tool for construction workers. *Journal of Construction Engineering and Management*, *145*(5), pp. 04019025. doi: 10.1061/(ASCE)CO.1943-7862.0001647.