Cambridge Journal of Regions, Economy and Society 2022, **15,** 663–682
https://doi.org/10.1093/cjres/rsac025
Advance Access publication 18 June 2022

Urban-regional disparities in mental health signals in Australia during the COVID-19 pandemic: a study via Twitter data and machine learning models

Siqin Wang[a], Mengxi Zhang[b], Xiao Huang[c], Tao Hu[d], Zhenlong Li[e], Qian Chayn Sun[f], Yan Liu[a]

[a]*School of Earth and Environmental Sciences, University of Queensland, Room 537, Chamberlain Building, St Lucia, Brisbane, QLD, 4076, Australia, S.W.: s.wang6@uq.edu.au; Y.L.: yan.liu@uq.edu.au*
[b]*Department of Nutrition and Health Science, Ball State University, Indiana, USA, mzhang2@bsu.edu*
[c]*Department of Geosciences, University of Arkansas, Arkansas, USA, xh010@uark.edu*
[d]*Department of Geography, Oklahoma State University, Oklahoma, 74078, USA, tao.hu@okstate.edu*
[e]*Geoinformation and Big Data Research Laboratory, Department of Geography, University of South Carolina, South Carolina, USA, zhenlong@sc.edu;*
[f]*School of Science, RMIT University, Melbourne, Victoria, Australia, chayn.sun@rmit.edu.au*

Received on August 27, 2021; editorial decision on May 20, 2022; accepted on May 30, 2022

This study establishes a novel empirical framework using machine learning techniques to measure the urban-regional disparity of the public's mental health signals in Australia during the pandemic, and to examine the interrelationships amongst mental health, demographic and socioeconomic profiles of neighbourhoods, health risks and healthcare access. Our results show that the public's mental health signals in capital cities were better than those in regional areas. The negative mental health signals in capital cities are associated with a lower level of income, more crowded living space, a lower level of healthcare availability and more difficulties in healthcare access.

Keywords: COVID-19, mental health, healthcare access, spatial disparity, Twitter, machine learning models
JEL classifications: C31, R50, R23, C23

Introduction

The COVID-19 pandemic and the policies implemented to control the spread of the virus have created stressful situations and challenges around the world, including the fear of contracting the virus, financial and employment losses, and government-mandated restrictions on movement and physical and social interactions (Betsch, 2020). These challenges are particularly detrimental to the mental health of individuals.

The effects of COVID-19 on people's mental health may vary across individuals with different demographic and socioeconomic status, and distinguishable across the urban and regional/rural space with distinct features in demographic and socioeconomic profiles, health risks, and access to healthcare facilities (Summers-Gabr, 2020; Sun and Lyu, 2020). Monitoring and measuring the urban-regional disparity of people's mental health status during the pandemic, and examining the potential factors contributing to such disparity, are crucial to provide evidence for place-based health planning policy intervention, as well as the provision of mental health services towards the post-pandemic era.

The current studies on mental health during the COVID-19 pandemic apply either survey-based assessments or advanced modelling techniques, such as machine learning algorithms (Balcombe, 2020). Survey-based studies conducted in different geographic contexts are inevitably subject to issues associated with the small data size, such as high cost, under-representativeness, and limited spatial and temporal coverage (e.g., Fisher et al., 2020; Newby et al. 2020; Van Rheenen et al., 2020). Most of the survey-based studies are limited to a certain period of the pandemic and focus on the early stage of the pandemic, which also lacks geographic information to reflect the spatial variation of mental health conditions. Another stream of research uses social media data and advanced modelling techniques to understand public's mental reactions to a range of COVID-related issues such as home schooling (Ewing and Vu, 2021), social restriction policies (Zhou et al., 2021) and vaccination (Hu et al., 2021). However, these studies are limited to certain states in Australia or relatively short time periods and also lack the coverage to understand the disparity between capital cities and regional areas. Subsequently, there is a pressing need to advance our understanding of people's mental health conditions across the urban and regional areas over the full temporal spectrum of the pandemic.

To fulfil the knowledge deficit, our study aims to investigate the urban-regional disparity of the public's mental health during the pandemic by addressing the following research questions: 1) How do the public's mental health signals vary across capital cities and regional areas? 2) How do the public's mental health signals change along the pandemic timeline? and 3) To what extent the public's mental health signals are associated with the demographic and socioeconomic profile of areas, their health risks and access to healthcare services? Drawing on 244,406 geotagged tweets in Australia from 1 January, 2020 to 31 May, 2021, we employed machine learning techniques to measure and classify the disparity in the public's mental health signals across the capital cities and regional areas in Australia. We further examined the interrelationships among mental health, the demographic and socioeconomic profiles of the neighbourhoods they live, their health risks and access to healthcare services, drawing on the demographic and socioeconomic data, health risk and healthcare access data retrieved from the Australian Urban Research Infrastructure Network (AURIN), 2020. Using the social media data with a large spatial and temporal coverage, our study contributes, for the first time, a nationwide examination of the public's mental health signals in Australia and the disparity between capital cities and their regional counterparts. We also demonstrate a novel empirical framework to systematically measure, classify, and map mental health signals nationwide, through which the role of public health policies and mental health services can be assessed in the wake of the global pandemic.

Background

Measuring mental health in the COVID-19 studies

Current studies on mental health relating to the COVID-19 pandemic are largely survey-based, focusing on certain social groups or populations in particular regions (Balcombe, 2020). Survey-based studies have been conducted with different sizes of study populations in different geographic contexts; however, they are inevitably subject to small-data issues such as the limited sample sizes and temporal coverages due to restricted financial resources (e.g., Fisher et al., 2020; Van Rheenen et al., 2020). These studies offer some valuable insights on factors that may increase a person's vulnerability to experiencing psychological distress such as depression, financial stress, health anxiety, contamination fears accompanied by the increasing level of alcohol use and decreasing level of physical activities (e.g., Newby et al., 2020; Van Rheenen et al., 2020). However, most existing studies focus on the early stage of the pandemic, with survey data also lacking geographic information to reflect the spatial variation of mental health signals.

Another set of research utilises crowdsourcing data, combined with the rapidly evolving computational infrastructures and intelligent algorithms (e.g., machine and/or deep learning) that offer exciting possibilities for monitoring both population-level and individual-level mental health status (Cotfas et al., 2021). In particular, social media data, as a well-established data source that has been applied in politics, businesses and disaster management, has been increasingly used in population health monitoring and other mental health applications (Conway and O'Connor, 2016). Qualitative or text-based data in social media (e.g., third-person pronouns and anger words) are the potential indicators of social media users' self-reported mental health problems (Coppersmith et al., 2014). The massive and insightful content portrayed and outlined by highly engaged social media users provide unprecedented opportunities for collective emotion and affective analysis (Liang et al., 2019). Based on the nature of social media data, a series of analyses via advanced machine learning and natural language processing techniques have been developed to monitor and track the public's mental health signals towards vaccination (Hu et al., 2021), social distancing policies, stay-at-home and lockdown orders (Zhou et al., 2021) and work-from-home requirements (Ewing and Vu, 2021) during the pandemic. There is great potential for using machine learning techniques within mental health studies, while there is an emerging critique that the effective application of machine learning is mediated by research design and bound up with a wide range of complex, interwoven challenges (Thieme et al., 2020). These challenges include generating large-scale, high-quality datasets to represent the diversity of the population (Bone et al., 2017), mitigating the obstacles (e.g., errors, uncertainty and bias) for the deployment of machine learning algorithms into real-world systems (Thornicroft et al., 2007), as well as considering far reaching personal, societal and economic implications in mental health contexts (Rudin et al., 2019). Despite of these challenges, machine learning techniques have offered new routes for learning patterns of human behaviour, identifying mental health symptoms and risk factors, and assisting in the detection, diagnosis and treatment of mental health problems. Considering that traditional survey methods are time and labour-consuming with limited spatial and temporal coverages, investigating the public's mental health during the COVID-19 pandemic over a longer timeline and with a larger spatial scale, and exploring the potential factors impacting on the disparity in mental health between urban and regional areas using public sourced social media data is needed.

Mental health and demographic and socioeconomic profiles of neighbourhoods

The current studies examining the prevalence of mental health problems have been largely influenced by Diez Roux's pathways theory (Diez-Roux, 1998). The pathways theory describes how the demographic and socioeconomic profile at the individual level and collectively at the aggregated population level (e.g., neighbourhoods) might contribute to the disparities of mental health across areas via individual and contextual pathways. The socioeconomic profile of neighbourhoods is commonly measured in several ways, including labour forces, income and financial status (e.g., mortgage stress), which may further affect travel and residential choices and dwelling types (e.g., living space, difficulty in moving or access to transport) (Meyer et al., 2014). These measures in socioeconomic profiles of neighbourhoods are particularly important in investigating mental health status during a pandemic. The common findings show that mental health problems (e.g., depression) are more likely to appear in low-socioeconomic neighbourhoods (Meyer et al., 2014). These studies also imply that the mental health problems that are associated with the pandemic appear more prevalent in low-socioeconomic areas with a concentrated low-income and unemployed population, or more essential workers who need to work onsite, as well as in areas with relatively crowded living space where the virus transmission is high (Zhang et al., 2021). In addition, the socioeconomic profile can be further compounded by the demographic profile of neighbourhoods, such as the age structure and race/ethnicity. For instance, the ageing neighbourhoods (e.g., retirement villages) and minority-concentrated suburbs (e.g., the Hispanic group in the U.S.) are subject to more severe mental health problems (e.g., worry and fear of infection), given the higher infection rate among the elderly and the Hispanic group comprising a higher proportion of essential workers compared to non-Hispanic groups (Penner et al., 2021). Given these considerations, we take onboard a set of indicators measuring the demographic and socioeconomic profiles of neighbourhoods as potential factors influencing the public's mental health.

Mental health and physical health

Another set of factors potentially associated with mental health are individuals' physical health status, such as chronic conditions (Talen and Mann, 2009). People with chronic diseases may experience negative emotions that further increase the probability of developing mental health issues. Such chronic diseases, including hypertension (Sparrenberger et al., 2009) and overweight and obesity (Talen and Mann, 2009), have been widely discussed in the current literature and observed to be associated with the onset of anxiety, stress and depression (Kretchy et al. 2014; Pan et al., 2015). Youth and adolescents who were overweight were more likely to report self-stigma, depression, anxiety and feelings of worthlessness (Chan et al., 2019), which may have long-lasting consequences on mental health. Furthermore, health risk behaviours such as alcohol consumption and smoking are potentially associated with various mental health issues, and poor mental health could be an enduring risk factor for heavy alcohol consumption (Shahab et al., 2014). The onset of these chronic conditions and health risk behaviours at the individual level assembles the prevalence of health risks at the aggregated population level, that is, the level of health risks in neighbourhoods. Subsequently, our study takes into account health risks in neighbourhoods when examining people's mental health during the pandemic.

Mental health and the spatial disparity of healthcare access

A study by Fisher et al. (2020) shows that around 25% of Australians had reported experiencing mild to moderate symptoms of depression or anxiety at the early stage of the COVID-19 pandemic. Lack of access to mental healthcare services and shortage of mental health providers may result in such mental health issues not being resolved properly or in a timely fashion (Lake and Turner, 2017). In Australia, the access to mental health services in rural/regional areas is considerably lower than that in major cities (Australian Institute of Health and Welfare (AIHW), 2019). In particular, people living in remote regional areas reported three times less access to mental health services subsidised by Medicare (i.e., Australia's universal health insurance scheme), compared to those who live in major cities (AIHW, 2019). The low rate of access to mental health services might be attributed to the limited number of mental health professionals and healthcare facilities in rural and remote Australia. Recent studies indicate an unequal distribution of mental health professionals (e.g., nurses and psychologists) across the metropolitan areas and regional/rural areas in Australia, with major cities having the most adequate workforce resource compared to other areas (Sutarsa et al., 2021). Apart from the availability/provision of healthcare services, access to mental health services is also influenced by other barriers that may access, such as the lack of connectivity to public transport services or the lack of private vehicles to drive to the healthcare facilities. Subsequently, people tend to use healthcare services and facilities that are within shorter travel distances or time from their homes more often (Ghorbanzadeh et al., 2020). Thus, it is critical to adjust for the availability of healthcare services and the level of difficulty in healthcare access in the examination of mental health across capital cities and regional areas.

In summary, through an empirical study in Australia, the key objective of our study is to provide nuanced but plausible insights into understanding the mental health of a nation and its urban-regional disparities, which is applicable in other countries aiming for better spatial justice and social harmony.

Study context, data and analysis

Study context

Australia is the largest developed country in the Southern Hemisphere, with a total population of nearly 26 million and a total area of around 7.61 million square kilometres (Australia Bureau of Statistics (ABS), 2020). Australia is highly urbanised, with over 80% of its population living in cities. The nation's capital city is Canberra, also known as the Australian Capital Territory (ACT), and the other states/territories are (the state capital cities are listed in brackets): New South Wales (Sydney), Victoria (Melbourne), Queensland (Brisbane), West Australia (Perth), South Australia (Adelaide), Tasmania (Hobart), and Northern Territory (Darwin). According to the Greater Capital City Statistical Area Structure (ABS, 2016), each state is divided into a greater capital city area and the remaining regional area. For instance, the State of New South Wales (NSW) is divided into the Greater Sydney Area and Rest of NSW. In this paper, we simplified the terminology using Sydney and Beyond Sydney to represent Greater Sydney and the Rest of NSW, respectively, in later analysis (Figure 1); such simplification also applied to other states/territories with the exception of ACT having only the capital city for the whole territory. Over 66% of the Australian population live in the greater capital city areas of the eight states/territories, with Sydney being the largest (with around 4.9 million population in 2016), followed by Melbourne (4.5 million) and Brisbane (2 million) (ABS, 2020). Our analysis first looked at the comparison between the capital cities aggregated as one unit and the regional areas, also aggregated as one unit, and then focused on the comparison between each capital city and its regional counterpart.

Data

We utilised Twitter academic full track application programming interface (AFT-API) to search and retrieve geotweets in Australia (Twitter, 2020). Compared to the normal Twitter API that returns 1% of the total tweets due to privacy concerns, AFT-API enables us to fully retrieve tweets with the pre-defined

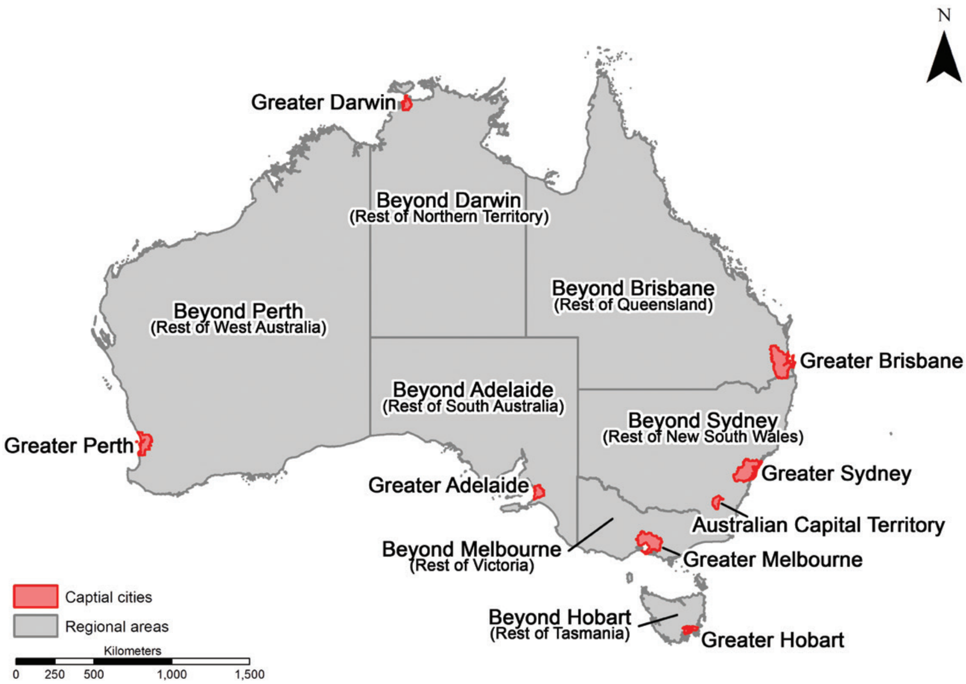


Figure 1. *Capital cities and regional areas in Australia*

queries, which improves the data coverage and representativeness (Twitter, 2020). We defined the searching terms as '*pandemic, epidemic, virus, covid*, and vaccine**'; the search timespan was defined as from 1 January, 2020 to 31 May, 2021 with the country defined as '*AU*' (Australia). Consequently, 244,406 geotweets were retrieved from the total of over 860 million tweets in Australia. These geotweets contain X, Y coordinates that were retrieved in two ways: 1) an accurate pair of X, Y coordinates of a place where a geotweet was posted if a user activated the positioning function in Twitter, and 2) an rough pair of X, Y coordinates converted from the name of a place (including cities and neighbourhood) if a user only indicated the name of that place in Twitter; such places were then assigned the X, Y coordinates as the centroid of that place. The geotweets were further aggregated to Statistical Area level 2 (SA2) as the second smallest unit in the Greater Capital City Statistical Area Structure (ABS, 2016) given that SA2 areas with an average population of about 10,000 persons serve as an appropriate unit compared to the full size of capital cities and the name of SA2 areas are identifiable. The SA2 areas with the number of tweets less than 17 were excluded, given it was assumed that there should be at least one tweet posted in each SA2 over the entire research period of 17 months (Jan 2020 to May 2021). We recognised that there were a disproportionate number of tweets concentrated in the SA2 area where the centroid of a capital city was located (not always the city centre) if users roughly tag a city in a tweet but they may reside in other SA2 areas. Thus, we further excluded eight SA2 areas containing the centroid of each capital city (e.g., the SA2 named 'Hughesdale' containing the centroid of Melbourne) to reduce the data bias. The spatial coverage of geotweets ranged from the highest of 69.9% in Melbourne to the lowest of 28.24% in Canberra, shown in the statistical summary and distribution of tweets provided in the supplementary material (Table S1 and Figure S1).

Demographic and socioeconomic data, and health risk and healthcare access data were retrieved from the AURIN online data portal (2020) at the level of Population Health Areas (PHA). PHA was developed by Public Health Information Development Unit (2016) and comprised a combination of whole SA2s and/or multiple segmented SA2s. We then used a correspondence file, as a cross-tabulated table containing the proportions of each SA2 falling into each PHAs), to convert the data in PHAs to SA2s in 2016, in order to join with the spatial boundary of SA2s and to align with the geotweets data aggregated at the SA2 level (ABS, 2016). The demographic and socioeconomic profile of one SA2 includes the measures of age structures, income, labour force, living space, needs for childcare, whether to have vehicles and be able to move. Health risk data include the proportion of people with poor health status, alcohol use, overweight, and high blood pressure in one SA2. Healthcare access data include the measures of hospital admissions and general practitioners (GP), and difficulty in healthcare access. More detailed definitions of these measures are provided in Table 1.

Analysis

Machine learning models to detect sentiment and emotion

We commenced with a series of data preprocessing to the geotweet records using Python 3.9.6 (details provided in 'Data preprocessing' section in the Supplementary materials). We then employed the Valence Aware Dictionary for sEntiment Reasoning (VADER) model to analyse the sentiment of each geotweet (VADER, 2021). The VADER model is a lexicon and rule-based sentiment analysis tool that has been specifically attuned to sentiments expressed in qualitative contexts (e.g., social media posts) (Hutto et al., 2014). VADER sentiment analysis relies on a machine learning algorithm and an open-source dictionary library that maps lexical features and heuristic expressions to emotion intensities known as sentiment scores (VADER, 2021). It returns four sentiment scores, including positive, negative, neutral, and compound scores. The first three scores are measured as a ratio of the number of words that fall in the respective categories (positive, negative and neutral sentiment) to the total number of words in each geotweet record, respectively. The compound score is a weighted composite score that is further generated based on the ratio of positive, negative and neutral sentiment. The value of a compound score ranges from -1 (most negative) to +1 (most positive). In this study, we used the compound score as the indicator of mental health given it is comparable across different geographic contexts (e.g., capital cities and regional areas). More details of VADER are provided in the work by Hutto et al. (2014).

To further interrogate the wide range of emotions that may be not reflected roughly by the positive and/or negative sentiment generated by VADER, we further used NRCLex, developed based on the National Research Council Canada Affect Lexicon and the Natural Language ToolKit (NLTK) library's WordNet synonym sets (Bird et al., 2009), containing approximately 27,000 words, to detect the emotional tendency of a given body of texts in geotweets. NRCLex differentiates the types of emotions via a word matching algorithm based on a documented affection dictionary and the association of the texts with four pairs of primary bipolar emotions (i.e., eight basic emotions): joy (feeling happy) versus sadness (feeling sad); anger (feeling angry) versus fear (feeling of being afraid); trust (stronger admiration and weaker acceptance) versus disgust (feeling something is wrong or nasty); and surprise (feeling unprepared for something) versus anticipation (looking forward positively to something). Among these eight types of emotions, fear, sadness, anger and

Table 1. *Name and definition of variables used in the regression modelling.*

Dependent variable	Definition
Binary variable representing a SA2 belonging to either the positive or negative group in mental health signals	1 (positive group): SA2s with a sentiment score higher than the average of all SA2s in the study area ** (this is the reference group) 0 (negative group):SA2s with a sentiment score lower than the average of all SA2s in the study area **
Independent variables	
Demographic and socioeconomic profiles	
Age 65 and over	% of people at and above 65 over the total population in one SA2
Aboriginal*	% of aboriginal population over the total population in one SA2
Unemployment	% of unemployed people over the total population in one SA2
Low-income	% of low-income households over the total number of households in one SA2
Mortgage stress	% of households under financial stress from mortgage or rent over the total number of households in one SA2
No vehicle*	% of households with no motor vehicle over the total number of households in one SA2
Crowded living*	% of people living in crowded dwellings over the total population in one SA2
Childcare need	% of households needed by childcare over the total number of households in one SA2
Difficulty in moving	% of people aged 18 years and over who encountered barriers to access to places (e.g. lack of public transport) over the total population in one SA2
Health risk	
Poor health	% of people aged 15 years and over with fair or poor self-assessed health over the total population in one SA2
Alcohol use	% of people aged 18 years and over who consumed more than two standard alcoholic drinks per day on average over the total population in one SA2
Overweight	% of people aged 18 years and over who were overweight over the total population in one SA2
High blood pressure	% of people aged 18 years and over who had high blood pressure over the total population in one SA2
Healthcare access	
Hospital and general practitioners (GP)	Total admissions of hospitals and GP numbers per 100 population in one SA2
Difficulty to access healthcare service	% of people aged 18 years and over who experienced a barrier to accessing healthcare when needed it in the last 12 months over the total population in one SA2

*This variable was log transformed to ensure the data is normally distributed.
**The study area is all regional areas in Australia for Model 1, all capital cities in Australia for Model 2, Greater Sydney metropolitan area for Model 3, Greater Melbourne metropolitan area for Model 4, and Greater Brisbane metropolitan area for Model 5.

disgust correspond to negative mental signals, while joy, anticipation, trust and surprise correspond to positive mental signals. The result of NRCLex was the number of words in each type of emotion in each geotweet record, which was further aggregated to a certain spatial unit (e.g., a capital city) and calculated as the proportion of words in each type of emotion over the total (named as percentages of emotion hereinafter).

Binary logistic regression

We then applied a binary logistic regression (BLR) (Hosmer et al., 1997) to examine how the mental health signals associate with the three sets of indicators measuring the demographic and socioeconomic profiles, health risks and healthcare access. SA2s with the compound sentiment score was transformed into a binary variable by comparing the sentiment score in each SA2 with the average in all SA2s of the study area, with 1 indicating higher-than-the-average SA2s (termed as the positive group; this group is set as the reference group in subsequent modelling) and 0 indicating lower-than-the-average SA2s (termed as the negative group). This binary variable is used as the dependent variable in the BLR, given that a binary dependent variable is most common in logistic regression which can produce a probability of a certain class, i.e., a SA2 being positive or negative in mental health signals. The BLR model is expressed as:

$$\text{Log}\left(\frac{P(Y=0)}{P(Y=1)}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \cdots + \beta_i x_i \quad (1)$$

where, $Y$is the dependent variable with two possible values, i.e., $Y = 1$ (the reference group) when the sentiment score in a SA2 is higher than the average sentiment score in all SA2s and $Y = 0$ otherwise; x_1 to x_i are the independent variables defined in Table 1; andβ_0 to $\beta_i$are the regression coefficients for each variable. The odds (o) indicating the likelihood of a SA2 belonging to a positive or negative SA2 is computed as:

$$o = 10^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \cdots + \beta_n x_n} \quad (2)$$

The corresponding probability (p) is calculated as:

$$p = \frac{o}{1 + o} \quad (3)$$

We ran a total of five BLR models. Models 1 and 2 are for regional areas and capital cities, respectively, and Models 3–5 are for Sydney, Melbourne and Brisbane as the three largest capital cities, respectively. Within each model, there are three sub-models (e.g., Models 1–1, 1–2, and 1–3), with the first sub-model involving the demographic and socioeconomic profile of neighbourhoods as the independent variables, the second also including the level of health risks, and the third further including healthcare access factor to the model. The significance levels were set at 0.1, 0.05 and 0.01.

Results

Change of mental health signals along the pandemic timeline

Figure 2 shows the temporal change of the public's mental health signals, illustrated by the sentiment scores in capital cities and regional areas. The sentiment score in capital cities is higher than that in regional areas over the whole timeline of the pandemic except for two short periods, one in January 2020 and one in March 2021. From February to March 2020, the sentiment scores in both capital cities and regional areas had an obvious increase and continued to increase afterwards to May 2020 but slightly decreased from May to July 2020. This corresponds to the second wave of the pandemic Australians experienced during this period. From September to November 2020, the sentiment scores in both capital cities and regional

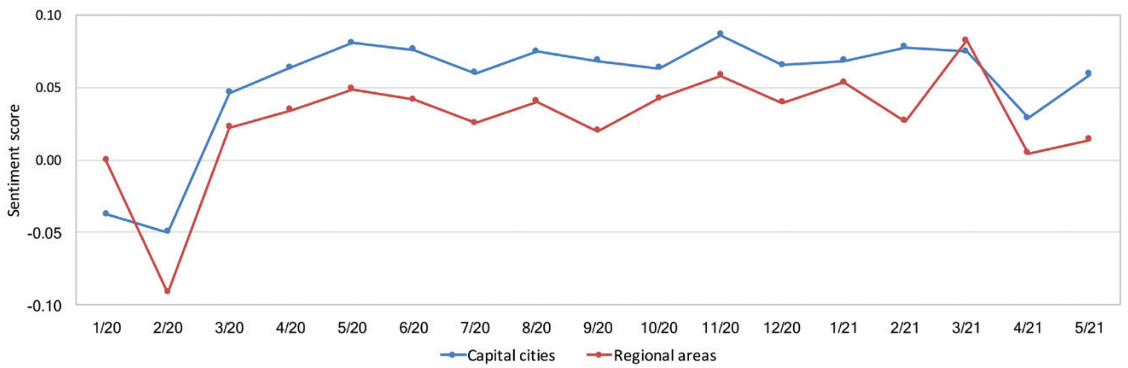


Figure 2. *Temporal change of the sentiment score in capital cities and regional areas*

areas increased to the peak in November 2020. From December 2020 to March 2021, the sentiment score in capital cities stayed relatively stable compared to that in regional areas experiencing more fluctuations. However, the sentiment scores in both capital cities and regional areas had an obvious decrease from March to April 2021. Overall, the public's mental health signals in capital cities were observed to be higher than that of their counterparts in the regional areas over the year-long period from February 2020 to February 2021.

The temporal change of the eight types of emotions provides a more detailed map of the variations of the public's mental health signals between the capital cities and the regional areas (Figure 3). At the early stage of the pandemic (January to March 2020), the feelings of fear had an obvious decrease in both the capital cities and regional areas while there were increases in trust and joy. From March 2020 to May 2021, the overall trend of these eight types of emotion in both capital cities and regional areas are relatively stable, and the feeling of fear accounts for the largest proportion of emotion (reflected by the highest position of the dark blue solid line), followed by trust, anticipation, sadness, joy, anger, surprise and disgust. The comparison of the percentages of emotion in capital cities (solid lines) and regional areas (dash lines) shows that there are no substantial differences in the percentages of emotion between capital cities and regional areas from March 2020 to May 2021, reflected by the small gap in a pair of solid and dash lines in one colour; while before March 2020, in regional areas there was a higher percentage in fear and a lower percentage in trust compared to capital cities. However, during the rest of the pandemic, the variations of emotion in between capital cities and regional areas is less clear, with the observation that the percentage in a certain type of emotion in capital cities is higher than that in regional areas in some months (e.g., April to July 2020 for fear) but lower in other months (e.g., April to May 2021 for fear).

Spatial disparity of mental health signals

We present the sentiment scores for each greater city area and their respective regional counterparts (e.g., Sydney and Beyond Sydney as we defined in Section 3.1), illustrating the subtle variations of the public's mental health signals of the country (Figure 4). Overall, the mean values of sentiment scores in all capital cities of the eight states/territories are higher than that of their regional counterparts, indicating that the public's mental health signals in cities tend to be more positive than those in regional areas. Moreover, the range of sentiment scores in one capital city (e.g., Sydney, Melbourne, Brisbane, and Perth) is wider than that in its counterpart (e.g., beyond Sydney, beyond Melbourne,

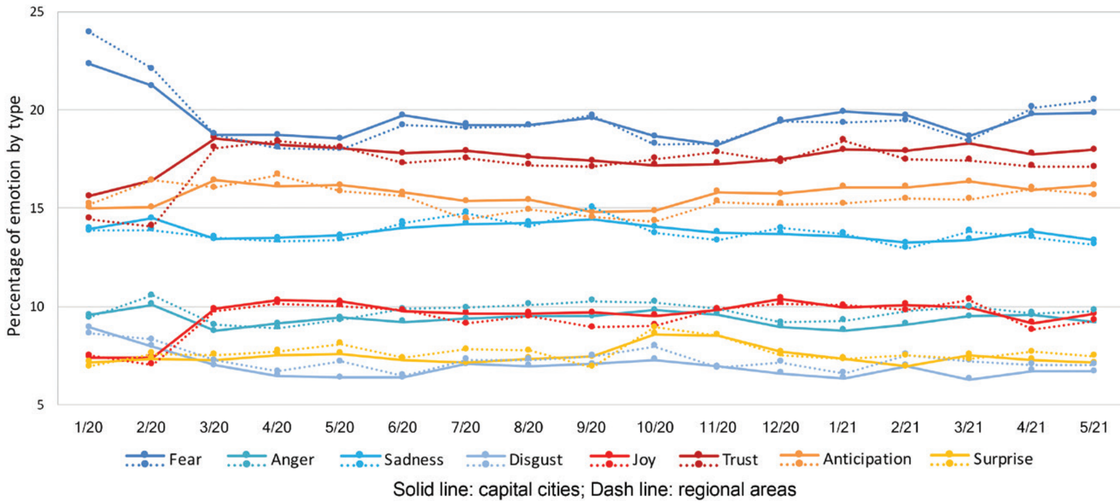
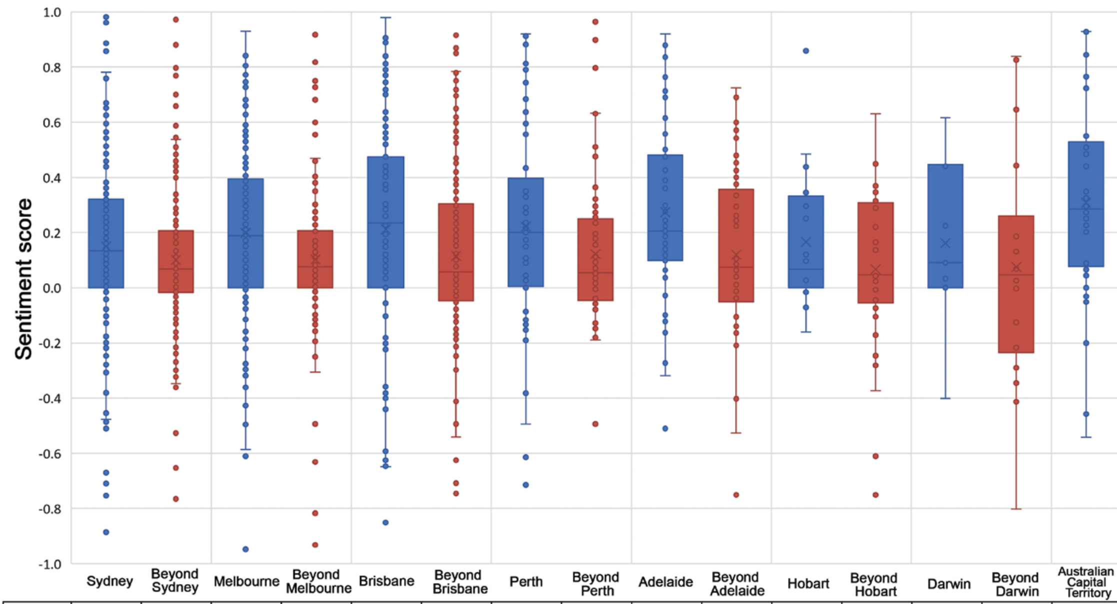


Figure 3. *Temporal change of emotion by type in capital cities and regional areas*



	Sydney	Beyond Sydney	Melbourne	Beyond Melbourne	Brisbane	Beyond Brisbane	Perth	Beyond Perth	Adelaide	Beyond Adelaide	Hobart	Beyond Hobart	Darwin	Beyond Darwin	Australian Capital Territory
Mean	0.150	0.101	0.199	0.103	0.212	0.115	0.221	0.122	0.274	0.120	0.166	0.067	0.162	0.075	0.308
Median	0.135	0.069	0.189	0.075	0.235	0.057	0.200	0.055	0.205	0.074	0.067	0.047	0.091	0.047	0.285
Std. Deviation	0.321	0.251	0.301	0.280	0.393	0.327	0.348	0.264	0.310	0.310	0.244	0.289	0.309	0.428	0.351
Minimum	-0.885	-0.765	-0.948	-0.932	-0.851	-0.745	-0.715	-0.494	-0.511	-0.751	-0.161	-0.751	-0.401	-0.802	-0.542
Maximum	0.981	0.972	0.929	0.917	0.980	0.915	0.920	0.964	0.920	0.725	0.859	0.630	0.617	0.839	0.929

Captial cities Regional areas

Figure 4. *Box-plots illustrating the variations of sentiment scores between capital cities and regional areas*

beyond Brisbane, and beyond Perth, respectively). In addition, the range of sentiment scores in the three largest state capital cities (i.e., Sydney, Melbourne, and Brisbane) is wider than that in the medium- and small-size capital cities (i.e., Perth, Adelaide, Hobart and Darwin). This indicates that although the mental health status of people in capital cities tends to be more positive than those in the regional areas, such signals are also more diverse and varied in cities than their regional counterparts.

Figure 5 reveals the variations of the emotions by type in the eight capital cities and their respective regional areas. In each of the graphs in Figure 5, a bar in a positive direction indicates a certain type of emotions (coloured differently) has a larger percentage in the city than that in its regional counterpart; the height of the stacked bar in each month reflects the extent the sum of the percentages for different types of emotion in the capital cities is higher or lower than that its counterpart of regional areas. In January and February 2020, all capital cities except Brisbane have higher percentages of negative emotion (including fear, anger, sadness and disgust, as illustrated in a blue colour theme in Figure 5) compared to the regional areas, shown by the stacked bar in the positive direction. In contrast, Brisbane has a higher percentage of positive emotion (including joy, trust, anticipation and surprise, as illustrated in a red–orange colour theme in Figure 5) compared to its counterpart of regional areas in January 2020. From February 2020 to May 2021, there are relatively minor variations in emotion between Melbourne and Brisbane and their regional areas compared to other capital cities. From January to April 2021, the percentages of positive emotion in Brisbane, Perth, Adelaide and Darwin are higher than their counterparts of regional areas, while the percentages of negative emotion in the regional areas beyond Sydney, Melbourne and Hobart are higher than the percentages in their capital cities. It is quite surprising to see that Melbourne, where the second wave of the pandemic in Australia was centered had higher percentages in positive emotion than its regional areas from June to October 2020, but shifted to higher percentages in negative emotion than its regional areas from November 2020 to March 2021. It is possibly due to the

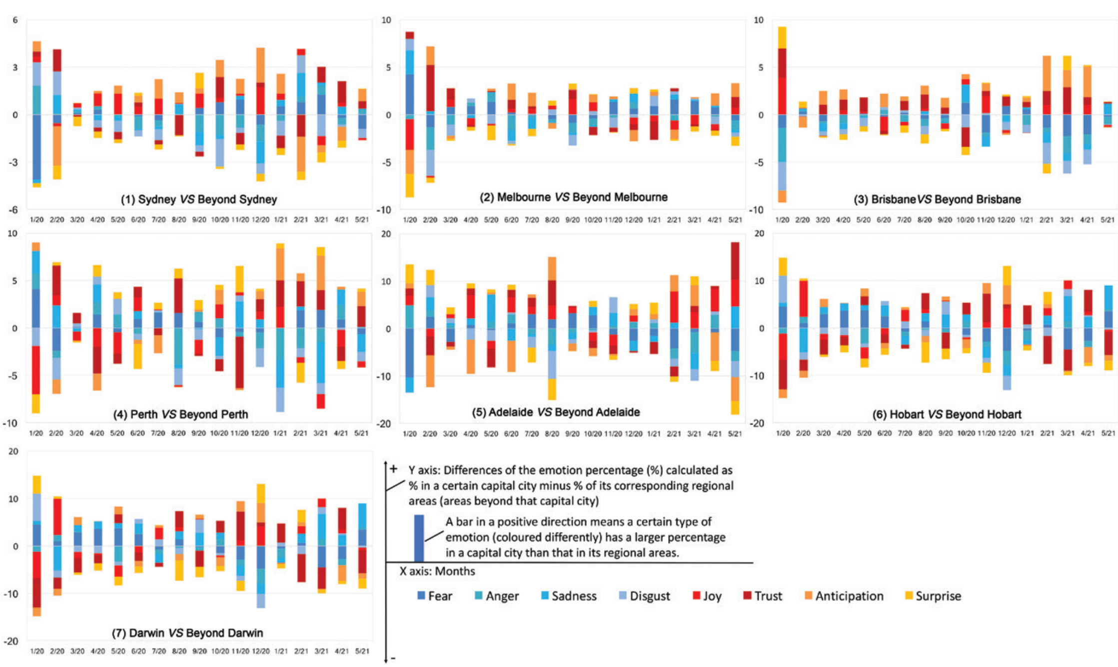


Figure 5. *Variations of the emotion types between capital cities and regional areas*

long-term implementation of social restriction policies (e.g., lockdown and home-dwelling orders) after the second wave of the pandemic from June to October 2020.

Relationships between mental health signals and the demographic and socioeconomic profiles of neighbourhoods, health risks and healthcare access

We further explored the extent to which the spatial and temporal variations of the mental health signals between capital cities and regional areas are associated with the demographic and socioeconomic profile of neighbourhoods, their health risks and healthcare access. In Model 1 (including sub-Models 1–1, 1–2, and 1–3 in Table 2), low-income has negative coefficients (at least $p < 0.1$, and with odds ratios smaller than 1) while the areas with crowded living and childcare need have positive coefficients (at least $p < 0.1$, correspondingly with odds ratios larger than 1). It means that people living in regional areas with lower income levels, more crowded living spaces, and higher needs for childcare are more likely to exhibit negative mental signals. The proportion of the aboriginal population is observed to have negative coefficients (at least $p < 0.1$, correspondingly with odds ratios smaller than 1) only in Model 1–2 and 1–3. By adding the health risks and healthcare access variables in Models 1–2 and 1–3, the Nagelkerke R^2 increased slightly, from 0.12 in Model 1–1 to 0.13 in Model 1–2 and 0.14 in Model 1–3; however, none of the health risks and healthcare access variables are significant at $p < 0.1$. This implies that the level of health risks and healthcare access in regional areas is not relevant to the mental health signals of people living in those regions.

For capital cities (Model 2, including Model 2–1, 2–2 and 2–3 in Table 3), the demographic and socioeconomic profiles of neighbourhoods that are significantly (at least $p < 0.1$) associated with mental health signals include age at and above 65 (a positive coefficient of 0.086 in Model 2–1 correspondingly the odds ratio larger than 1), aboriginal (the odds ratio smaller than 1 in Model 2–3), unemployment and low-income (all odds ratios smaller than 1), crowded living (odds ratios larger than 1 in Model 2–1 and 2–3), and difficulty in moving (all odds ratios larger than 1). It means that people living in areas with a smaller proportion of aboriginal population, more crowded living space, and more difficulty in moving and transport tend to exhibit negative mental signals. The additional involvement of health risk variables in Model 2–2 increases the Nagelkerke R^2 from 0.21 in Model 2–1 to 0.33 in Model 2–2. All health risk and healthcare access variables are observed to be significantly (at least $p < 0.05$) associated with mental health signals. More specifically, poor health, alcohol use and high blood pressure have positive coefficients (odds ratios larger than 1); while being overweight has negative coefficients (odds ratios smaller than 1). It reveals that people living in the areas with a higher proportion of the population having issues in poor health, alcohol use and high blood pressure tend to be more likely to display negative mental signals. Further adding in healthcare access variables in Model 2–3 increases the Nagelkerke R^2 from 0.33 in Model 2–2 to 0.48 in Model 2–3 and both healthcare access variables are significantly ($p < 0.01$) associated with mental health signals. More specifically, the number of hospital admissions and GPs has an odds ratio of 0.598 [0.391, 0.915] while difficulty to healthcare access has an odds ratio of 1.047 [1.032, 1.062]. It means that people residing in areas with lower availability of hospitals and GPs and more difficulties in healthcare access tend to be more likely to have negative mental signals. This finding in capital cities is distinct from that in regional areas, reflecting the spatial disparity of health risks and healthcare access in capital cities and regional areas that is further related to the public's mental health status.

Table 2. *Regression results for regional areas.*

Regional areas (beyond capital cities)	Model 1–1		Model 1–2		Model 1–3	
	Coefficient	Odds ratio [A] (95% CI)	Coefficient	Odds ratio [A] (95% CI)	Coefficient	Odds ratio [A] (95% CI)
Demographic and socioeconomic profile						
Age 65 and over	0.059	1.061 (0.985, 1.143)	0.056	1.058 (0.979, 1.143)	0.049	1.051 (0.971, 1.137)
Aboriginal	-0.062	0.939 (0.864, 1.021)	-0.083*	0.921 (0.833,1.018)	-0.082*	0.921 (0.833, 1.018)
Unemployment	0.101*	1.106 (0.974, 1.256)	0.100	1.106 (0.961, 1.272)	0.098	1.102 (0.959, 1.267)
Low-income	-0.037**	0.964 (0.929, 1.000)	-0.041**	0.960 (0.921, 1.001)	-0.039*	0.962 (0.922, 1.003)
Mortgage stress	-0.029	0.971 (0.928, 1.016)	-0.029	0.972 (0.925, 1.020)	-0.037	0.964 (0.914, 1.017)
No vehicles	0.016	1.016 (0.949, 1.088)	0.022	1.022 (0.950, 1.100)	0.028	1.028 (0.954, 1.108)
Crowded living	0.185**	1.204 (1.041, 1.392)	0.212**	1.236 (1.046, 1.460)	0.207**	1.230 (1.042, 1.453)
Childcare need	0.062*	1.064 (0.990, 1.143)	0.074*	1.077 (0.996, 1.165)	0.071*	1.074 (0.992, 1.162)
Difficulty in moving	-0.039	0.962 (0.690, 1.343)	-0.105	0.900 (0.593, 1.368)	-0.145	0.865 (0.563, 1.329)
Health risk						
Poor health			0.032	1.033 (0.920, 1.160)	0.019	1.019 (0.903, 1.150)
Alcohol use			0.025	1.026 (0.949, 1.108)	0.024	1.024 (0.948, 1.106)
Overweight			-0.010	0.990 (0.876, 1.119)	0.002	1.002 (0.883, 1.136)
High blood pressure			-0.021	0.979 (0.782, 1.226)	-0.029	0.972 (0.777, 1.216)
Healthcare access						
Hospital and GP					-0.004	0.996 (0.986, 1.007)
Difficulty in healthcare access					0.177	1.193 (0.738, 1.930)
Nagelkerke R^2	0.12		0.13		0.14	
OCP	65.92		67.41		67.16	
Number of SA2	660		660		660	

Note: $^*p < 0.1$; $^{**}p < 0.05$; $^{**}p < 0.01$.
The reference group is the areas concentrated by people with relatively positive mental status.
CI, confidence interval; *OCP*, an overall correct percentage of prediction.
[A]Numbers in the bracket under each odd radio is the range of confidence interval of that odd radio at the level of 95%.

Table 3. *Regression results for capital cities.*

All eight capital cities	Model 2–1		Model 2–2		Model 2–3	
	Coefficient	Odds ratio [A] (95% CI)	Coefficient	Odds ratio [A] (95% CI)	Coefficient	Odds ratio [A] (95% CI)
Demographic and socioeconomic profile						
Age 65 and over	0.086***	1.09 (1.021, 1.164)	0.599	1.82 (1.334, 2.483)	0.036	1.037 (0.954, 1.126)
Aboriginal	0.011	1.011 (0.968, 1.056)	0.078	1.081 (1.005, 1.163)	-0.039**	0.962 (0.904, 1.024)
Unemployment	-0.210***	0.811 (0.717, 0.917)	-0.042*	0.959 (0.901, 1.02)	-0.126*	0.882 (0.759, 1.025)
Low-income	-0.026*	0.975 (0.941, 1.01)	-0.100	0.905 (0.784, 1.043)	-0.004**	0.996 (0.951, 1.044)
Mortgage stress	0.019	1.019 (0.978, 1.062)	-0.019	0.981 (0.94, 1.023)	0.033	1.033 (0.981, 1.088)
No vehicles	0.022	1.022 (0.955, 1.093)	0.012	1.012 (0.968, 1.059)	-0.014	0.986 (0.909, 1.071)
Crowded living	0.041**	1.042 (0.979, 1.109)	-0.006	0.994 (0.921, 1.074)	0.092***	1.097 (1.014, 1.186)
Childcare need	0.035	1.036 (0.963,1.114)	0.105***	1.111 (1.026, 1.202)	-0.040**	0.961 (0.882, 1.047)
Difficulty in moving	0.599***	1.82 (1.334, 2.483)	0.005**	1.005 (0.928, 1.089)	0.268***	1.308 (0.832, 2.055)
Health risk						
Poor health			0.219***	-0.803 (-0.909, -0.71)	0.112**	1.118 (1.019, 1.226)
Alcohol use			0.078**	1.081 (1.001, 1.168)	0.072*	1.074 (0.991, 1.164)
Overweight			-0.292***	0.747 (0.652, 0.856)	-0.198***	0.820 (0.709, 0.95)
High blood pressure			0.488***	1.629 (1.294, 2.05)	0.304**	1.355 (1.064, 1.724)
Healthcare access						
Hospital and GP					-0.513**	0.598 (0.391, 0.915)
Difficulty in healthcare access					0.046***	1.047 (1.032, 1.062)
Nagelkerke R^2	0.21		0.33		0.48	
OCP	66.40		71.70		74.5	
Number of SA2	705		705		705	

Note: *$p < 0.1$; **$p < 0.05$; **$p < 0.01$.
The reference group is the areas concentrated by people with relatively positive mental status.
CI, confidence interval; *OCP*, an overall correct percentage of prediction.
[A]Numbers in the bracket under each odd radio is the range of confidence interval of that odd radio at the level of 95%.

We also examined the extent the public's mental health is associated with the demographic and socioeconomic profile of neighbourhoods, their health risks and healthcare access in the three largest capital cities, including Sydney, Melbourne and Brisbane (Table S2–4). The results show that the social and spatial disparity in demographic and socioeconomic profile across areas in capital cities are potentially associated with the public's mental health. More specifically, low-income (with odds ratios larger than 1) and crowded living space (with odds ratios smaller than 1) are observed to be associated with mental health signals in Sydney and Melbourne; while difficulty in moving and transport (with odds ratios larger than 1) is associated with mental health signals only in Sydney and mortgage stress (with odds ratios larger than 1) is associated with mental health signals only in Brisbane. However, health risks and healthcare access variables are observed to be not significantly associated with mental health signals in these three capital cities.

Discussion and conclusion

Drawing on social media data and machine learning techniques, this study makes an initial attempt to reveal the urban-regional disparity of the public's mental health status during a global scale pandemic and examine the extent to which mental health signals are associated with the demographic and socioeconomic profiles of neighbourhoods, their health risks and healthcare access. This has been under-explored in current scholarship in the Australian context. Urban and regional/rural areas are featured by different demographic and socioeconomic characteristics and configured by various healthcare facilities and services, which can affect people's sentimental and emotional reactions to the pandemic. It is especially the case for capital cities where people with negative mental health signals are observed to live in areas with limited healthcare resources and more difficulty in healthcare access. Our finding align with the observations in previous studies focusing on the mental health of the general population in different geographic contexts. More specifically, the minority-concentrated neighbourhoods (e.g., with a higher proportion of aboriginal population) have a higher likelihood of negative mental health signals; this finding echoes the findings in the US that Hispanic/Latinx groups are subject to greater risks for worsened mental health due to the pandemic (Penner et al., 2021). Socioeconomic disparity across neighbourhoods with different levels of income, mortgage stress and living space are subject to the discrepancy of mental health signals; such an observation is also found in case studies in China (Liu et al., 2021), US (McKnight-Eily et al., 2021), Caribbean regions (Llibre-Guerra et al., 2020), the United Kingdom (Pierce et al., 2020) and European countries (Reznik et al., 2020). Moreover, neighbourhoods with the concentrations of the population at higher health risks are subject to a higher likelihood of negative mental signals, reflecting that the onset of health risks (e.g., poor health status, high blood pressure and drinking habits) may worsen mental health in the face of the pandemic. This finding is less observed in COVID-19 related studies; however, more prevalent in mental health studies before the pandemic (e.g., Hardy et al., 2013; Kwan et al., 2016). Furthermore, another disparity between capital cities and regional areas observed in our study is the availability of healthcare facilities and services and the difficulty in healthcare access which is further associated with the prevalence of negative mental health signals. Easy access to mental health facilities could help persons with mental illness to integrate into the patient community and facilitate peer-led interventions to improve medical self-management (Druss et al., 2010). It has been observed that older men living in rural Australia may have a similar incidence of mental health problems compared to older women, despite a lower rate of diagnosis (Fitzpatrick et al., 2021). The obstacles to men obtaining mental health treatment in rural

Australia include the reluctancy of emotional expression, non-disclosure of distress and difficulties in seeking and getting help (Kennedy et al., 2021). Thus, neighbourhoods configured by better healthcare facilities serve as a contextual pathway to protect against the onset of psychological problems, improving the perception of local residents in feeling safe, secure and protected and thus reducing mental disorders.

Our findings provide evidence for policy making and implementation in public health. In response to the urban-regional disparity of the public's mental health signals, regional health authorities and government agencies should distribute more healthcare services in these vulnerable areas prior to any unprecedented events; intervention programs and mental health services should be introduced or put in place in regional areas and urban areas in capital cities which are currently less covered by healthcare facilities. National level assistance, such as digital mental health guidelines, are also recommended in regional areas with a higher proportion of socioeconomically disadvantaged populations. Mental health strategies should be varied along pandemic event timelines, with monitoring and more effective decision-making processes to allow prompt responses to the rapidly changing pandemic. Our models and research framework may be useful for governments and policymakers at various levels to monitor public's mental health signals in future pandemics and public health crises.

This study has several limitations that need to be noted and can be improved in future studies. First, the geotweet data used in our study is non-random sample, since the number of Twitter users is dependant on population density and thus tend to concentrate in densely populated areas (e.g., capital cities) whilst are less prevalently in regional/remote areas, which may bring bias to the analytical results (Zhou et al., 2021). Second, Twitter users are typically younger, avid users of social media apps and the Internet (Huang et al., 2020), while older individuals are less likely to use smartphones and Aboriginal Australians have lower internet accessibility and are relatively inactive in social media (Povey et al., 2016). Thus, the current geotweets may be skewed to represent the opinions and perceptions of the subsection of the population. In addition, the current geotweets are in English, with non-English ones excluded from sampling. Future studies using multilingual tweets might better represent different ethnicities and spoken languages. Third, our analysis did not involve the COVID-19 measures (e.g., confirmed cases, and mortality and morbidity rates) given such data is not available at the SA2 level. The spread and severity of COVID-19 infection have been reported to associated with the public's mental status (Kenerly et al., 2022), thus we encourage future studies to take on board these COVID-19 measures to enhance the analytical robustness. Fourth, the built environment in the capital cities and regional areas can be very different (e.g., the density, diversity, and design of neighbourhoods); the measures of the built environment need to be included in the regression modelling to better reveal the context pathway through which urban/rural space affects the public's mental health during the pandemic. In addition, it has been reported that the type and stringency of policy implementation over different phases of the pandemic also affects people's overall sentiment (Zhou et al., 2021). Thus, the measures of policies can be included in future modelling work.

While our study provides empirical evidence of using machine learning techniques in mental health studies, there remain difficulties and challenges in translating the analytical findings to the real-world treatment of at-risk cohorts. The empirical framework developed in our study can be improved by integrating with highly scalable and accurate machine learning platforms (e.g., symbiotic marching learning systems) and trials of local system dynamics models which have been used in monitoring the suicidal behaviours of Australian youth

(Iorfino et al., 2021). Future work should also give attentions to the usability challenges for machine learning (Zytek et al., 2021). It requires the researchers to have sufficient skills and contextual knowledge in mental health and medicine to develop human-centred research procedure, reliable interpretation of modelling results, and cross validation of findings based on different approaches (e.g., traditional survey-based assessment, and clinic protocol and trials) (Thieme et al., 2020).

To conclude, our study contributes a novel empirical framework using social media and machine learning techniques to systematically classify and measure the urban-regional disparity of mental health signals of a nation. Our approach is designed in a manner that can readily be augmented into an ongoing monitoring capacity and extended to other nations. Our findings in the interrelationship among mental health, the demographic and socioeconomic profiles of neighbourhoods, their health risk and healthcare access provide important evidence for the smart deployment of finite mental health services and place-based health planning towards the post-pandemic period and beyond.

Acknowledgments

This study is supported by the NCRIS-enabled Australian Urban Research Infrastructure Network (AURIN) with the project name—Integrated Heat Vulnerability Assessment Toolkit for Australian Cities, AURIN, High Impact Projects 2021.

Conflict of interest statement

None.

References

Australia Bureau of Statistics. (2016) 1270.0.55.001—Australian Statistical Geography Standard (ASGS): volume 1—main structure and greater capital city statistical areas. Available at: https://www.abs.gov.au/ausstats/abs@.nsf/mf/1270.0.55.001 [accessed 18 Aug 2021].

Australia Bureau of Statistics. (2020) National, state and territory population. Available at: https://www.abs.gov.au/statistics/people/population/national-state-and-territory-population/latest-release [Accessed 18 Aug 2021].

Australian Institute of Health and Welfare. (2019) Rural and remote health. Available at: https://www.aihw.gov.au/reports/australias-health/rural-and-remote-health [Accessed 18 Aug 2021].

Australian Urban Research Infrastructure Network. (2020) AURIN data. Available at: https://aurin.org.au/resources/data/ [Accessed 18 Aug 2021].

Balcombe L., and De Leo D. (2020) An integrated blueprint for digital mental health services amidst COVID-19. *JMIR Mental Health*, **7**: e21718.

Betsch C. (2020) How behavioural science data helps mitigate the COVID-19 crisis. *Nature Human Behaviour*, **4**: 438–438.

Bird, S., Loper, E., and Klein, E. (2009) Natural Language Processing with Python. Available at: https://www.nltk.org/ [Accessed 18 Aug 2021].

Bone D., Lee C. C., Chaspari T., Gibson J., and Narayanan S. (2017) Signal processing and machine learning for mental health research and clinical applications [perspectives]. *IEEE Signal Processing Magazine*, **34**: 196–195.

Chan K. L., Lee C. S., and Cheng C. M., et al (2019) Investigating the relationship between weight-related self-stigma and mental health for overweight/obese children in Hong Kong. *The Journal of Nervous and Mental Disease*, **207**: 637–641.

Conway M., and O'Connor D. (2016) Social media, big data, and mental health: current advances and ethical implications. *Current Opinion in Psychology*, **9**: 77–82.

Coppersmith, G., Harman, C., and Dredze, M. (2014) Measuring post traumatic stress disorder in Twitter. In Eighth International AAAI Conference on Weblogs and Social Media (pp. 579–582).

Cotfas L. A., Delcea C., Roxin I., Ioanăş C., Gherai D. S., and Tajariol F. (2021) The longest month: analyzing COVID-19 vaccination opinions dynamics from tweets in the month following the first vaccine announcement. *IEEE Access*, **9**: 33203–33223.

Diez-Roux A. V. (1998) Bringing context back into epidemiology: variables and fallacies in multilevel analysis. *American Journal of Public Health*, **88**: 216–22.

Druss B. G., Zhao L., and Silke A. et al (2010) The Health and Recovery Peer (HARP) Program: a peer-led intervention to improve medical self-management for persons with serious mental illness. *Schizophrenia Research*, **118**: 264–270.

Ewing L. A., and Vu H. Q. (2021) Navigating ‘Home Schooling’during COVID-19: Australian public response on Twitter. *Media International Australia*, **178**: 77–86.
Fisher J. R., Tran T. D., and Hammarberg K. et al (2020) Mental health of people in Australia in the first month of COVID-19 restrictions: a national survey. *Medical Journal of Australia*, **213**: 458–464.
Fitzpatrick S. J., Read D., Brew B. K., and Perkins D. (2021) A sociological autopsy lens on older adult suicide in rural Australia: addressing health, psychosocial factors and care practices at the intersection of policies and institutions. *Social Science and Medicine*, **284**: 114196.
Ghorbanzadeh M., Kim K., Ozguven E. E., and Horner M. W. (2020) A comparative analysis of transportation-based accessibility to mental health services. *Transportation Research Part D Transport and Environment*, **81**: 102278.
Hardy S. A., Francis S. W., Zamboanga B. L., Kim S. Y., Anderson S. G., and Forthun L. F. (2013) The roles of identity formation and moral identity in college student mental health, health-risk behaviors, and psychological well-being. *Journal of Clinical Psychology*, **69**: 364–382.
Hosmer D. W., Hosmer T., Le Cessie S., and Lemeshow S. (1997) A comparison of goodness-of-fit tests for the logistic regression model. *Statistics in Medicine*, **16**: 965–80.
Hu T., Wang S., and Luo W., et al (2021) Revealing public opinion towards COVID-19 vaccines with twitter data in the United States: spatiotemporal perspective. *Journal of Medical Internet Research*, **23**: e30854.
Huang X., Li Z., Jiang Y., Li X., and Porter D. (2020) Twitter reveals human mobility dynamics during the COVID-19 pandemic. *PLoS One*, **15**: e0241957.
Hutto C., and Gilbert E. (2014) Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the International AAAI Conference on Web and Social Media (Vol. **8**, No. 1).
Iorfino F., Occhipinti J. A., and Skinner A. et al (2021) The impact of technology-enabled care coordination in a complex mental health system: a local system dynamics model. *Journal of Medical Internet Research*, **23**: e25331.
Kenerly M. J., Shah P., and Patel H. et al (2022) Altered mental status is an independent predictor of mortality in hospitalized COVID-19 patients. *Irish Journal of Medical Science*, **191**: 21–26.
Kennedy A., Cosgrave C., Macdonald J., Gunn K., Dietrich T., and Brumby S. (2021) Translating co-design from face-to-face to online: an Australian primary producer project conducted during COVID-19. *International Journal of Environmental Research and Public Health*, **18**: 4147.
Kretchy I. A., Owusu-Daaku F. T., and Danquah S. A. (2014) Mental health in hypertension: assessing symptoms of anxiety, depression and stress on anti-hypertensive medication adherence. *International Journal of Mental Health Systems*, **8**: 1–6.
Kwan M. Y., Arbour-Nicitopoulos K. P., Duku E., and Faulkner G. (2016) Patterns of multiple health risk–behaviours in university students and their association with mental health: application of latent class analysis. *Health Promotion and Chronic Disease Prevention in Canada rRsearch, Policy and Practice*, **36**: 163–170.
Lake J., and Turner M. S. (2017) Urgent need for improved mental health care and a more collaborative model of care. *The Permanente Journal*, **21**: 17–024.
Liang Y., Zheng X., and Zeng D. D. (2019) A survey on big data-driven digital phenotyping of mental health. *Information Fusion*, **52**: 290–307.
Liu L., Xue P., Li S. X., Zhang J., Zhou J., and Zhang W. (2021) Urban-rural disparities in mental health problems related to COVID-19 in China. *General Hospital Psychiatry*, **69**: 119–120.
Llibre-Guerra J. J., Jiménez-Velázquez I. Z., Llibre-Rodriguez J. J., and Acosta D. (2020) The impact of COVID-19 on mental health in the Hispanic Caribbean region. *International Psychogeriatrics*, **32**: 1143–1146.
McKnight-Eily L. R., Okoro C. A., and Strine T. W. et al (2021) Racial and ethnic disparities in the prevalence of stress and worry, mental health conditions, and increased substance use among adults during the COVID-19 pandemic—United States, April and May 2020. *Morbidity and Mortality Weekly Report*, **70**: 162–66.
Meyer O. L., Castro-Schilo L., and Aguilar-Gaxiola S. (2014) Determinants of mental health and self-rated health: a model of socioeconomic status, neighbourhood safety, and physical activity. *American Journal of Public Health*, **104**: 1734–1741.
Newby J. M., O’Moore K., Tang S., Christensen H., and Faasse K. (2020) Acute mental health responses during the COVID-19 pandemic in Australia. *PLoS One*, **15**: e0236562.
Pan Y., Cai W., Cheng Q., Dong W., An T., and Yan J. (2015) Association between anxiety and hypertension: a systematic review and meta-analysis of epidemiological studies. *Neuropsychiatric Disease and Treatment*, **11**: 1121–30.
Penner F., Ortiz J. H., and Sharp C. (2021) Change in youth mental health during the COVID-19

pandemic in a majority Hispanic/Latinx US sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, **60**: 513–523.

Pierce M., Hope H., and Ford T. et al (2020) Mental health before and during the COVID-19 pandemic: a longitudinal probability sample survey of the UK population. *The Lancet Psychiatry*, **7**: 883–892.

Povey J., Mills P. P. J. R., and Dingwall K. M. et al (2016) Acceptability of mental health apps for Aboriginal and Torres Strait Islander Australians: a qualitative study. *Journal of Medical Internet Research*, **18**: e5314.

Public Health Information Development Unit. (2016) Population health areas: overview. Available at: https://phidu.torrens.edu.au/about-phidu [Accessed 18 Aug 2021].

Reznik A., Gritsenko V., Konstantinov V., Khamenka N., and Isralowitz R. (2020) COVID-19 fear in Eastern Europe: validation of the fear of COVID-19 scale. *International Journal of Mental Health and Addiction*, **19**: 1903–1908. doi:10.1007/s11469-020-00283-3.

Rudin C. (2019) Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, **1**: 206–215.

Shahab L., Andrew S., and West R. (2014) Changes in prevalence of depression and anxiety following smoking cessation: results from an international cohort study (ATTEMPT). *Psychological Medicine*, **44**: 127–41.

Sparrenberger F., Cichelero F., and Ascoli A., et al (2009) Does psychosocial stress cause hypertension? A systematic review of observational studies. *Journal of Human Hypertension*, **23**: 12–19.

Summers-Gabr N. M. (2020) Rural–urban mental health disparities in the United States during COVID-19. *Psychological Trauma Theory, Research, Practice, and Policy*, **12**: S222–24.

Sun J., and Lyu S. (2020) Social participation and urban-rural disparity in mental health among older adults in China. *Journal of Affective Disorders*, **274**: 399–404.

Sutarsa N., Banfield M., Passioura J., Konings P., and Moore M. (2021) Spatial inequities of mental health nurses in rural and remote Australia. *International Journal of Mental Health Nursing*, **30**: 167–176.

Talen M. R., and Mann M. M. (2009) Obesity and mental health. *Primary Care: Clinics in Office Practice*, **36**: 287–305.

Thieme A., Belgrave D., and Doherty G. (2020) Machine learning in mental health: a systematic review of the HCI literature to support the development of effective and implementable ML systems. *ACM Transactions on Computer–Human Interaction (TOCHI)*, **27**: 1–53.

Thornicroft G., Rose D., Kassam A., and Sartorius N. (2007) Stigma: ignorance, prejudice or discrimination?. *The British Journal of Psychiatry*, **190**: 192–193.

Twitter. Twitter API—Academic Research product track. (2020) Available at:https://developer.twitter.com/en/products/twitter-api/academic-research [Accessed 18 Aug 2021].

VADER. (2021) VADER library. Available at: https://github.com/cjhutto/vaderSentiment [Accessed 18 Aug 2021].

Van Rheenen T. E., Meyer D., and Neill E., et al (2020) Mental health status of individuals with a mood-disorder during the COVID-19 pandemic in Australia: initial results from the COLLATE project. *Journal of Affective Disorders*, **275**: 69–77.

Zhang M., Wang S., and Hu T. et al (2021) Human mobility and COVID-19 transmission: a systematic review and future directions. *Annals of GIS*, 1–14. doi:10.1080/19475683.2022.2041725.

Zhou J., Yang S., Xiao C., and Chen F. (2021) Examination of community sentiment dynamics due to COVID-19 pandemic: a case study from a state in Australia. *SN Computer Science*, **2**: 1–11.

Zhou J., Zogan H., Yang S., Jameel S., Xu G., and Chen F. (2021) Detecting community depression dynamics due to COVID-19 pandemic in Australia. *IEEE Transactions on Computational Social Systems*, **8**: 1–10. Available online at: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9325873.

Zytek A., Liu D., Vaithianathan R., and Veeramachaneni K. (2021) Sibyl: understanding and addressing the usability challenges of machine learning in high-stakes decision making. *IEEE Transactions on Visualization and Computer Graphics*, **28**: 1161–71.