https://doi.org/10.1038/s44220-023-00147-4

Challenges and future directions for investigating the effects of urbanicity on mental health

Gunter Schumann

Check for updates

This Comment proposes to increase knowledge of the effects of urbanicity on brain and mental health by linking existing human spatial data with macroenvironmental and regional socioeconomic data. It introduces hypothesis-free models to capture the data and model life in the city and suggests refinements for future studies into conditions that will soon affect the majority of the earth's population.

We are all exposed to a complex environment that affects our health. Urbanicity is a pervasive environmental condition of global relevance. Urbanicity integrates diverse macroenvironmental factors, including land use, socio-economic and cultural factors, climate, pollution and others[1]. In 1950, less than 30% of the world's population lived in cities; this number has increased to currently 55% and is expected to rise to 68% in 2050[2]. The urban environment is associated with mental disorders, such as depression, anxiety and schizophrenia.

Complex, high-dimensional models are necessary to capture, disentangle and quantify the relative contributions of the various macroenvironmental influences present in urban life[3] and to assess their interactions, including with the microenvironment. Until recently, it has not been possible to establish such high-dimensional models that aim to recreate 'urban living', not least because of the lack of large datasets with sufficient statistical power. Now, however, the availability of large datasets, advanced research strategies and novel statistical models make this approach feasible[4].

Such high-dimensional models represent a major methodological shift to hypothesis-free approaches, which were first used in mass-univariate genome-wide association studies. They are a departure from traditional experimental approaches in life sciences, which are characterized by a narrow focus and highly controlled conditions. In these traditional approaches, knowledge has been assembled from fragmented components of empirical information, with less regard to the relations between the individual components. This approach may be appropriate for some questions in natural sciences, but is not best suited to capture the complex conditions affecting human health. Owing to its integration of diverse environmental factors and its relevance for human health, the study of urbanicity is a useful paradigm to develop analytical models that integrate multiple data modalities in order to capture realistic living environments with increasing fidelity.

Enriching existing datasets

To maximize the yield of investigations measuring the relation of urban living environments to brain, behaviour and mental health, enrichment of existing datasets as well as careful study planning and preparation are required (Fig. 1). In our experience, in the IMAGEN (https://www.imagen-project.org), STRATIFY (https://www.stratify-project.org), cVEDA (https://www.cveda-project.org) and environMENTAL (https://www.environmental-project.org) projects, several points proved critical.

Geoposition. Most existing studies and cohorts will not have the comprehensive characterization of the environmental factors necessary for studying urbanicity. Thus, enrichment of cohorts is critical. In the case of urbanicity or other macroenvironmental factors, external datasets can be linked via geoposition to individual participants, thus relating brain and behavioral measures to external environmental data, including satellite data[5] and pollution, climate or regional socioeconomic conditions[1]. To measure exposure over a lifetime, biographies of geoposition, typically limited to addresses of home, school or workplaces, can prove useful[5]. This information is easily obtained in many historical datasets, and in the environMENTAL consortium, we propose that care should be taken to include a comprehensive set of geopositions in the planning of new studies, as they can greatly enhance the environmental characterization at minimal cost and effort.

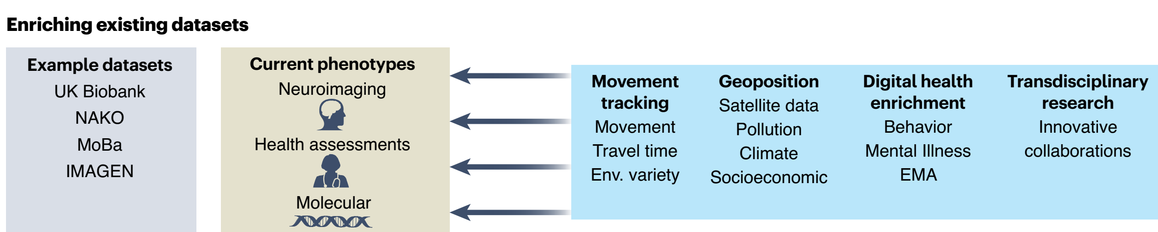

Fig. 1 | Proposed pathways towards enriching existing data sources with macro- and microenvironmental data to improve knowledge of the effects of urbanicity on mental health. Collected data will be used to generate hypothesis-free models to capture the high-dimensional complexity of environmental exposure enabling the creation of a quantifiable life-like model of life in the city. MoBa, Norwegian Mother, Father and Child Cohort Study.

Movement tracking. The quality of the geoposition determines the precision of the measure of environmental exposure. So far, we have simply used an anonymized form of the home address. However, it is not known how much time study participants actually spend at home, where else they spend their time and how they get to their destinations, all of which could influence environmental exposure and consequently their brain and behavior. Thus movement-tracking data, obtained under strict adherence to privacy requirements – for example, from passive registration of smartphones – might greatly increase precision of geoposition data. Alternatively, smartphone applications – such as the digital health app Streetmind – enable participants to decide when to switch on movement tracking, thus reducing data load and enabling annotation of movement paths (https://go.nature.com/3ZOfCdC).

Digital health enrichment. Most large-scale, population-based cohorts, such as UK Biobank (https://www.ukbiobank.ac.uk) and the German National Cohort (https://nako.de) cover a wide range of physical and mental health areas and thus have limited depth of behavioral assessments. To enhance mental health characterization, digital health tools are an option. Although it may not be easy to accommodate such internal enrichment in large cohorts, the gain can be considerable. For example, in the Streetmind app, such assessments involve self-rating instruments and participation in app-based ecological momentary assessments. Aside from an in-depth mental health assessment, this enrichment enables investigations of interactions between macro- and micro-environmental factors to identify subgroups who are at particular risk for mental illness in an urban environment. Together with movement data, digital health measures can be of use for mental health assessments as well for studies of physical health, such as cardiovascular or respiratory investigations.

Transdisciplinary research. Research on urbanicity and mental health can benefit from innovative collaborations; for example, with urban planners. Although the semantic gap between the disciplines can be considerable, there is important knowledge potential from such collaborations. For example, urban planners have developed measures to estimate occupancy and interior space using passive sensing technologies[6] and models to predict areas within a city that are prone to excessive heat[7]. Applying such refined models to existing mental health datasets will aid in describing more precisely the relation between urban-environmental factors, brain and behavior.

Study planning

As proposed above, the great statistical power of large population-based cohorts together with enrichment enable comprehensive, multi-modal modeling to elucidate the relation of an urban living environment with brain, behavior and mental health. Still, limitations remain, particularly if specific underlying biological, cognitive or psychological mechanisms for which data are not available in the large cohorts selected are to be investigated. Such limitations can be overcome by careful, stepwise study design. The guiding principle is the use of large population-based cohorts to identify features that reduce dimensionality. These features enable further investigation in dedicated datasets of much smaller size that are comprehensively characterized to address the research question of interest. This approach is bi-directional: findings that have been generated in specialized studies – for example, endophenotypes identified from clinical samples of schizophrenia – can be tested for generalization and moderation by environmental determinants of urbanicity at the population level.

The dynamics of urbanization are most pronounced in Asia, Latin America and Africa, whereas most cities in Europe, the USA, Japan and Australia are growing less rapidly or even decreasing in size. Thus, our understanding of the effects of urbanization on brain and behavior will remain incomplete, unless a global perspective is adopted. Comparative analyses, which are already challenging owing to different social, demographic, geographic, cultural and socio-economic conditions, are further compounded by varying conditions for data access, different assessment instruments and a general lack of harmonization across cohorts. Therefore, data federation and uniform acquisition of new measures of urbanicity are of critical importance[8]. To facilitate global comparisons, we have recently developed standardized satellite-based measures of urbanicity that are applicable globally[5]. Federated data analysis tools such as COINSTAC[9] enable simultaneous analyses of distributed datasets while maximizing patient confidentiality and respecting local data access regulations.

Refining the models

There are various multivariate methods to integrate different environmental factors that are relevant for urbanicity and link them to brain and behavior, including regression analyses, different machine learning approaches, deep learning techniques, and others. However, in our experience, these methods do not fully account for the effects of urbanicity because they do not optimize for best correlation between environmental factors with brain and behavioral symptoms. For this reason, we apply canonical correlation analyses, adjusted for sparsity and the number of data planes investigated[3,5,10]. Using this approach, we found that variance of behavioral symptom groups explained by urban-environmental variables in sparse canonical correlation analyses was more than ten times the variance of either principal components analyses or symptom scores of psychiatric disorders[3], suggesting that studies measuring the contribution of urban-environmental factors to psychiatric disorders that use methods not optimised for correlation may underestimate the relevance of environmental factors for mental health.

All the methods mentioned above rely on group-level statistics, which limit their use for individual predictions. Normative modeling is a framework for understanding differences at the level of a single subject or observation while mapping these differences in relation to a reference model. It has long been used in diverse applications, including growth charting in pediatric studies, neurocognitive tests and for calculating test score percentiles[11]. Normative modeling approaches have recently been adapted to clinical neuroscience and are available in the Predictive Clinical Neuroscience toolkit (https://pcntoolkit.readthedocs.io/en). Once stable relations between urban-environmental factors, brain and behavior are identified, the prediction of individual risk using normative models will become a highly relevant research aim.

There is an inherent assumption that urban living may cause mental health problems, but there are also reports suggesting that individuals who are vulnerable to mental disorders are more likely to move to socioeconomically deprived urban areas. It has also been proposed that unmeasured familial factors may account for the association between urbanicity and mental health[12]. Given that most cohort studies are currently limited to cross-sectional observations, the identification of causality remains a key challenge. Statistical models, such as mediation analyses and Bayesian causal-network analyses can provide assumptions of directionality, definitive confirmation can only be attained in interventional longitudinal analyses, which are not a realistic proposition in the study of urbanicity and mental health. Analyses of observational longitudinal datasets may be a pragmatic way to provide more robust evidence about the directionality of effects.

Outlook

Mental health has always been subject to macro-environmental conditions created by humans, but the number of people exposed to urban environments is unprecedented. Equally unprecedented is our ability to measure the different environmental factors that contribute to the various forms of urban living. To fully leverage this opportunity, we need to develop and refine conceptual and methodological approaches that integrate the available information and model urban living environments with increasing fidelity. Investigating the relation of the urban living environment to brain, behavior and mental health carries great promise for effective prevention and targeted early interventions, which will benefit millions.

Gunter Schumann [1,2] ✉

[1]Centre for Population Neuroscience and Stratified Medicine, Department of Psychiatry and Neuroscience, CCM, Charite University Medicine Berlin, Berlin, Germany. [2]Centre for Population Neuroscience and Stratified Medicine, Institute for Science and Technology of Brain-Inspired Intelligence, Fudan University, Shanghai, P. R. China.

✉e-mail: gunter.schumann@charite.de

Published online: 6 November 2023

References

1. Polemiti, E. et al. *Mol. Psychiatry* (in press).
2. Heilig, G. K. http://esa.un.org/wpp/ppt/paa/PAA_2012_Heilig.pdf (2012).
3. Xu, J. et al. *Nat. Med.* **29**, 1456–1467 (2023).
4. Schumann, G. et al. *JAMA Psychiatry* **80**, 1066–1074 (2023).
5. Xu, J. et al. *Nat. Hum. Behav.* **6**, 279–293 (2022).
6. Duygu Tekler, Z. & Chong, A. *Build. Environ.* **226**, 109689 (2022).
7. Huang, J. et al. *Energy Build.* **207**, 109580 (2020).
8. Schumann, G. et al. *Lancet Glob. Health* **7**, e32 (2019).
9. Gazula, H. et al. *Neuroinformatics.* **21**, 287–301 (2023).
10. Ing, A. et al. *Nat. Hum. Behav.* **3**, 1306–1318 (2019).
11. Rutherford, S. et al. *Nat. Protoc.* **17**, 1711–1734 (2022).
12. Evans, B. E. et al. *Nat Med.* **29**, 1322–1323 (2023).

Acknowledgements

The author is writing this Comment on behalf of the environMENTAL consortium. This work has received funding from the European Union. Complementary funding was received by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee (10041392 and 10038599). Views and opinions expressed are, however, those of the author only and do not necessarily reflect those of the European Union or European Health and Digital Executive Agency (HADEA). Neither the European Union nor HADEA can be held responsible for them.

Competing interests

The author declares no competing interests.