Editorial

Big Brain Data Initiative in Universiti Sains Malaysia: Challenges in Brain Mapping for Malaysia

Sharifah Aida Shekh Ibrahim[1,3], Nurfaten Hamzah[1,3], Athirah Raihanah Abdul Wahab[1,3], Jafri Malin Abdullah[1,2,3], Nurul Hashimah Ahamed Hassain Malim[3,4], Putra Sumari[3,4], Zamzuri Idris[1,2,3], Ariffin Marzuki Mokhtar[5,6], Ab Rahman Izaini Ghani[1,2,3], Sanihah Abdul Halim[3,7], Salmi Ab Razak[3,8]

[1] Department of Neurosciences, School of Medical Sciences, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia

[2] Department of Neurosciences, Hospital Universiti Sains Malaysia, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia

[3] Brain and Behaviour Cluster, School of Medical Sciences, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia

[4] School of Computer Sciences, Universiti Sains Malaysia, Pulau Pinang, Malaysia

[5] Department of Anesthesiology, School of Medical Sciences, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia

[6] Department of Anesthesiology, Hospital Universiti Sains Malaysia, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia

[7] Unit of Neurology, Department of Medicine, School of Medical Sciences, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia

[8] Unit of Paediatric Neurology, Department of Paediatrics, School of Medical Sciences, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia

Abstract

Universiti Sains Malaysia has started the Big Brain Data Initiative project since the last two years as brain mapping techniques have proven to be important in understanding the molecular, cellular and functional mechanisms of the brain. This Big Brain Data Initiative can be a platform for neurophysicians and neurosurgeons, psychiatrists, psychologists, cognitive neuroscientists, neurotechnologists and other researchers to improve brain mapping techniques. Data collection from a cohort of multiracial population in Malaysia is important for present

1

and future research and finding cure for neurological and mental illness. Malaysia is one of the participant of the Global Brain Consortium (GBC) supported by the World Health Organization. This project is a part of its contribution via the third GBC goal which is influencing the policy process within and between high-income countries and low- and middle-income countries, such as pathways for fair data-sharing of multi-modal imaging data, starting with electroencephalographic data.

Keywords: Big Data, Big Brain Data, brain mapping, neurosciences, initiative, neuroimaging

Introduction

The Fourth Industrial Revolution loads our life with tons of data (1). These data that require routine collection, storage, processing and analysis (2). The support of technological advances, it has led to the concept of 'Big Data' that describe data that is large and massive (1).

An Overview of Big Data Definition

The needs of Big Data analytics is steadily increasing in various field as it can be used in remodel operational processes of an organisation. The term of Big Data refers to a collection of large volume of both structured and unstructured data for the purpose of identifying unknown relationships or other informative attributes by using computational analysis that often involving advanced machine learning algorithms (3). It is an information assets that is often characterised by three properties which are high volume, velocity and variety. It requires specific technologies and analytical methods for its transformation into meaningful values (4). The challenges when dealing with Big Data include capture, curation, storage, search, sharing, transfer, analysis and visualisation (5). Although Big Data does not associate to any specific volume of data, Big Data storage often involve terabytes, petabytes and even exabytes of data captured over time (6).

In health research, there is a specific definition of Big Data proposed by the Health Directorate of the Directorate-General for Research and Innovation of the European Commission. The definition Big Data in health encompasses high volume, high diversity biological, clinical, environmental and lifestyle information collected from single individuals to large cohorts, in relation to their health and wellness status, at one or several time points (2).

The Importance of Big Data in Brain, Mind and the Neurosciences

Information has been the key to build better organisations and new developments. The organisation can optimally organise themselves to deliver the best outcomes if they have more information. That is why data collection is a significant part for every organisation (1). In healthcare, Big Data can be used in the prediction of diseases outcome, prevention of co-morbidities, mortality, premature deaths and disease development, improve treatment and quality of life and reduce the cost of medical treatment (7).

Today, patients want to participate in their health decision-making about their healthcare options or choices. This Big Data will help patients with providing up-to-date information to assist them to make the best decision and to comply with the medical treatment.

In brain, mind and neurosciences, Big Brain Data can lead to important discoveries; particularly, it can help us to understand the brain's structure and function, identify new biomarkers of brain pathology and increase the performance of neurotechnological devices (such as brain-computer interfaces [BCIs]) (3). This growing Big Data would also be of much use for the advanced machine learning algorithms, specifically artificial neural networks for deep learning and related methods.

Big Brain Data Initiative in Universiti Sains Malaysia with the Global Brain Consortium

Big Brain Data refers to the recording, collection and analysis of an individual brain data on a large scale by using neurotechnologies, like electroencephalogram (EEG), magnetoencephalography (MEG), near-infrared spectroscopy, magnetic resonance imaging (MRI) and functional magnetic resonance imaging (fMRI) (3).

The Big Data project on retrospective and prospective data for data hospital (pilot neuro) in Universiti Sains Malaysia (USM) was started on the 14th August 2017 (8). It was led by the Brain and Behaviour Cluster, School of Medical Sciences (previously known as Centre for Neuroscience Service and Research) in collaboration with the School of Computer Sciences with three objectives. The main objective was to establish a centralised clinical data storage platform for future references and research. The other two objectives are to create digital (paperless) access and data availability that can be accessed at university level by students, lecturers, clinicians and researchers.

This project was developed to contribute to the third Global Brain Consortium (GBC) goals which was influencing the policy process within and between high-income countries and low- and middle-income countries, such as pathways for fair data-sharing of multi-modal imaging data, starting with EEG (9). GBC is a global collaboration of brain, mind and neuroscientists across borders and disciplines to build a fluid and connected global research community that is better able to advance equitable solutions to priority health challenges worldwide (10). Malaysia was honoured to be invited to attend the GBC Plenary Meeting 2019 that was held on 9th May to 10th May 2019 at Montreal Neurological Institute-Hospital, Canada supported World Health Organization (WHO). Malaysia was represented by Professor Dato' Dr Jafri Malin Abdullah, the Chairman of Brain and Behaviour Cluster, School of Medical Sciences, USM. Continuation from that, he attended the online GBC Workshop 2020 from 27th February to 28th February 2020 at Varadero, Cuba in order to improve the international standards of the USM Big Data Project as shown in Figures 1 to 5.

The retrospective of MEGs, EEGs and MRIs were collected from year 2001 to 2018 from Hospital USM by two researchers, Sharifah Aida Shekh Ibrahim and Nurfaten Hamzah, from Health Campus, USM, meanwhile the interface and the storage platform was developed and design by Associate Professor Dr Nurul Hashimah Ahamed Hassain Malim, a lecturer from the School of Computer Sciences, USM. Currently, the project is on the testing phase.

Figure 6 shows the data pre-processing (curation) pipeline that had been established to understand the formats of data and the degree of aggregation needed from the various sources of input. Meanwhile, Table 1 shows the latest progress on Big Brain Data Initiative of USM.

Benefits of Big Brain Data to Researchers/Clinicians in Malaysia and the World

There are several benefits to researchers/ clinicians of Big Brain Data such as for our better understanding of the brain's anatomy including structure and it's function, identifying new biomarkers of brain pathology, neurotechnological devices performance improvement for example BCIs (3) and opportunity to explore and implementing of 'internet of things' (IOT) in health industries (1).
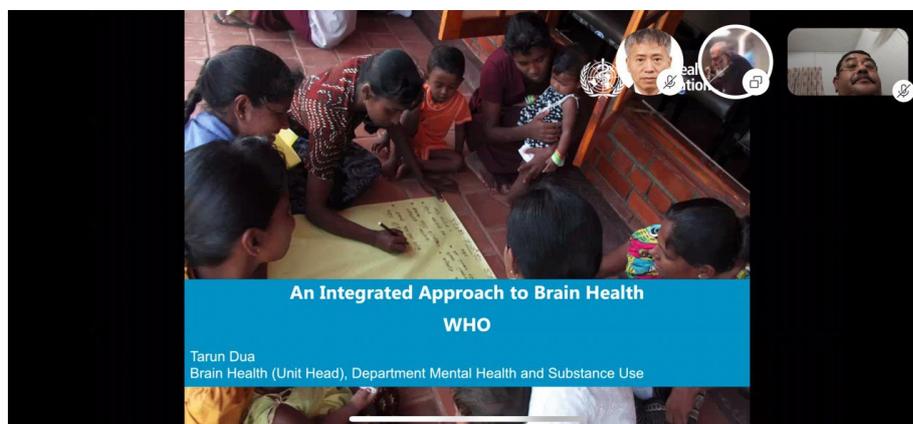

Figure 1.   Professor Dato' Dr Jafri Malin Abdullah joined the GBC Workshop 2020—online web conferencing. The topic 'An Integrated Approach to Brain Health' was presented by Dr Tarun Dua, Brain Health (Unit Head), Department Mental Health and Substance Use, WHO
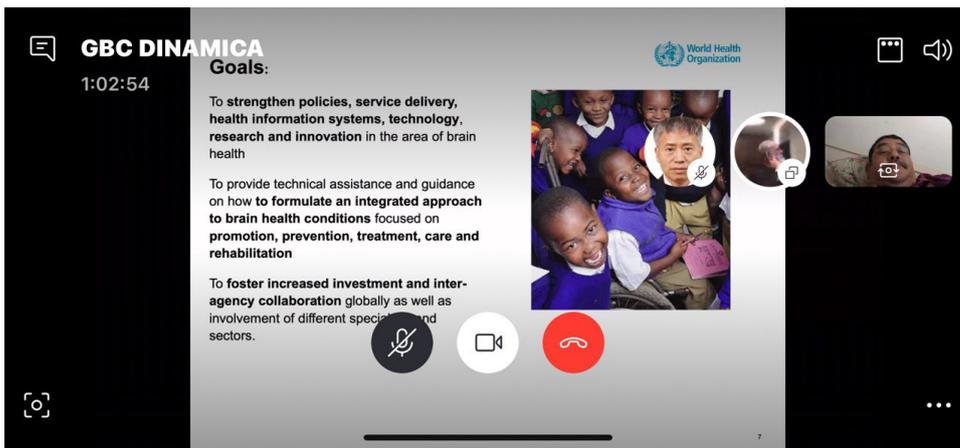
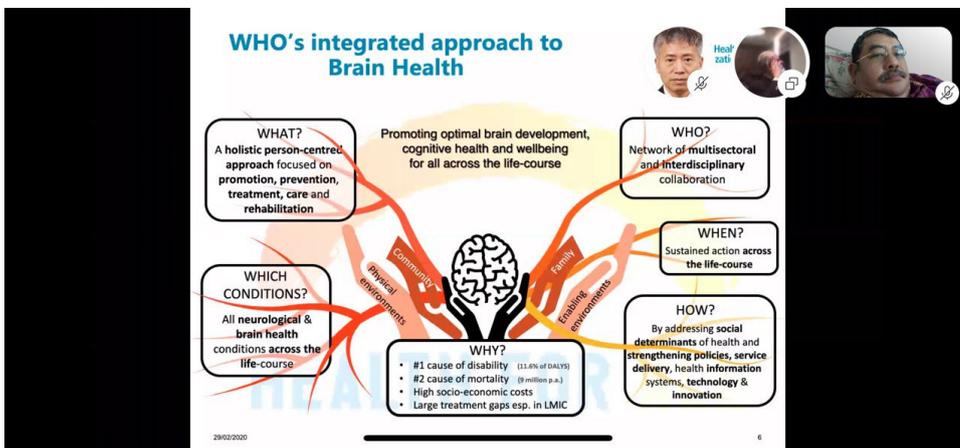**Figure 2.** Dr Tarun Dua presented slides on the three goals of an Integrated Approach to Brain Health

Figure 3. Promoting optimal brain development, cognitive health and well-being for all across the life-course

Figure 4. Workgroup 6 of the GBC 2020 presented validation from patients' advocates and other stakeholders (regulatory agencies, public health, pharma companies) and test from EEG paradigms, clinical applications and validation topic
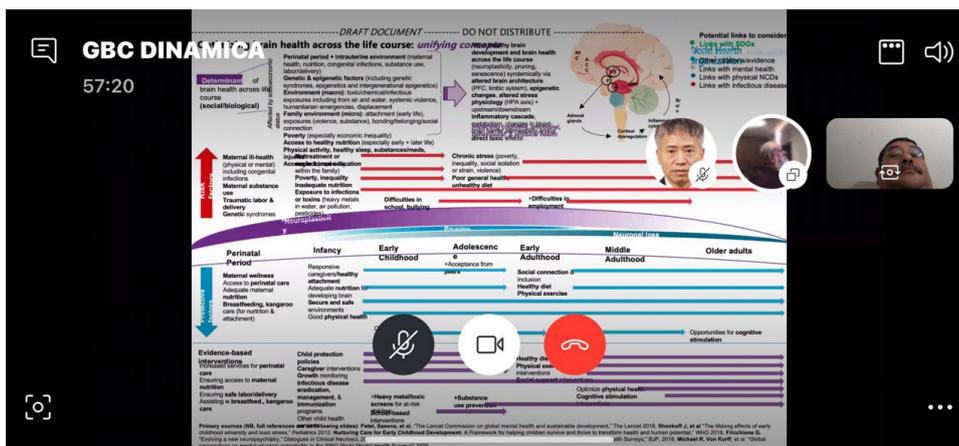
Figure 5. Optimising brain health across the life course: unifying concepts in terms of determinat of brain health across the course (social/biological), risk factors and protective factors
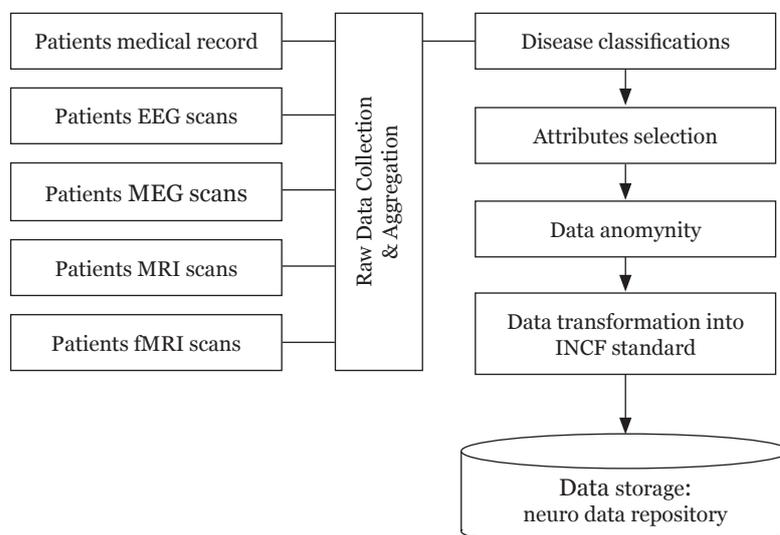

Figure 6. The data pre-processing (curation) pipeline © USM (8) starts with collection of records (medical records and scanning records) and followed by collection and aggregation, disease classifications, attribute selection, data anonymity, data transformation into international neuroinformatics coordinating facility standard and neuro data repository

Table 1. Raw data collection and aggregation of EEG and MEG are 100% collected. While MRI data collection is 15% because the number and size of the data is huge. The fMRI data collection is ongoing

	Raw data collection and aggregation (%)	Record date range	Disease classification (%)	Attributes selection (%)	Data anonymity (%)	Data transformation (%)	Number of records
EEG	100	2001–March 2018	50	80	100	2	4214
MEG	100	2008–February 2018	60	90	100	30	190
MRI	15	2002–October 2018	50	60	10	1	8876
fMRI	0	NA	0	0	0	0	19

An Overview of Brain Mapping

Brain mapping techniques have proven to be vital in understanding the molecular, cellular and functional mechanisms of the brain. Normal anatomical imaging can provide structural information on certain abnormalities in the brain. However there are many neurological disorders for which only structure studies are not sufficient. In such cases it is required to investigate the functional organisation of the brain. Brain mapping techniques can help in deriving useful and important information on these issues (11).

According to the Society for Brain Mapping and Therapeutics (SBMT), United States of America, brain mapping is clearly defined as the study of the structural and functional of the brain and spinal cord through the use of imaging. Mapping brain structural and functional connections through the whole brain is important for understanding brain mechanisms and the physiological bases of brain diseases (12). The development of high-resolution neuroimaging and multielectrode electrophysiological recording can be considered as a part of brain mapping that can help brain, mind and neuroscientists to map what the brain looks like as disease progress or as treatments work. The commonly neurotechnologies used are positron emission tomography and fMRI along with EEG, electrocorticography, MEG and, most recently, optical imaging with near-infrared spectroscopy (13).

The Challenges in Brain Mapping

Four common challenges in brain mapping has been identified which are scale, complexity, speed and integration.

i) Scale – large numbers of neurons and synapses, human brain simulation would push exascale computers to work beyond limit.

ii) Complexity – almost limitless set of parameters are require to produce a biologically faithful simulation of the brain. Brain's extracellular interactions and molecular-scale processes such as receptor binding are examples of details which not incorporated into simulation models.

iii) Speed – currently there is no technology can support large-scale simulations faster than in real time, normally models run slower. It is because human's brain processes such as development and learning took years or decades to perform.

iv) Integration – combination of smaller models of brain regions are needed in order to model functions that involve brain-wide networks. Top-down and bottom-up models must be integrated, for example, those that cast the brain as a hypothesis testing system and biophysical model that represent simulations (14–19).

Conclusion

The Big Brain Data Initiative in USM with objectives to establish a centralised clinical data storage platform for future references and research, creating digital (paperless) access and data availability which can be access at university level by students, lecturers, clinicians and researchers is reaching to the testing phase.

Moving to the next stage is to implement the brain mapping techniques for understanding the molecular, cellular and functional mechanism of the brain. Four commons challenges will be faced by researchers are scale, complexity, speed and integration. The brain-techniques and computer technologies should play their roles simultaneously in order to overcome the challenges.

Acknowledgements

Conflict of Interest

None.

Funds

Authors' Contributions

Conception and design: SASI, NH, JMA, NHAHM, ARAW, PS
Drafting of the article: SASI, NH
Critical revision of the article for important intellectual content: JMA, NHAHM, PS, ZI, AMM, ARIG, SAR, SAH
Final approval of the article: JMA, NHAHM, PS
Obtaining of funding: NHAHM
Administrative, technical, or logistic support: SASI, NH, NHAHM, ARAW, PS
Collection and assembly of data: SASI, NH

Correspondence

Professor Dato' Dr Jafri Malin Abdullah
MD (USM), PhD (University of Ghent, Belgium), FRCS (Ed), FACS (USA), DSCN (Belgium), FAMM, FRSM (UK), FICS (USA)

i)   Brain and Behaviour Cluster, School of Medical Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia.

ii) Department of Neurosciences, School of Medical Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia.

iii) Department of Neurosciences, Hospital Universiti Sains Malaysia, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia.

Tel: +609 7676300
Fax: +609 7673833
E-mail: brainsciences@gmail.com

References

1. Dash S, Shakyawar S, Sharma M, Kaushik S. Big Data in healthcare: management, analysis and future prospects. *J Big Data*. 2019;**6(1)**:54. https://doi.org/10.1186/s40537-019-0217-0

2. Pastorino R, De Vito C, Migliara G, Glocker K, Binenbaum I, Ricciardi W, et al. Benefits and challenges of Big Data in healthcare: an overview of the European initiatives. *Euro J Public Health*. 2019;**29(3)**:23–27. https://doi.org/10.1093/eurpub/ckz168

3. Kellmeyer P. Big Brain Data: on the responsible use of Brain Data from clinical and consumer-directed neurotechnological devices. *Neuroethics*. 2018. https://doi.org/10.1007/s12152-018-9371-x

4. De Mauro A, Greco M, Grimaldi M. A formal definition of Big Data based on its essential features. *Library Review*. 2016;**65(3)**:122–135. https://doi.org/10.1108/LR-06-2015-0061

5. Kumar D, Duvvuri BKK, Umar S. A study on Big Data and its importance. *Int J Appl Eng Res*. 2014;**9(20)**:973–4562. Available at: https://www.researchgate.net/publication/267626895_A_Study_on_Big_Data_and_its_Importance

6. Pence, H. What is Big Data and why is it important? *Journal of Educational Technology Systems*. 2014;**43(2)**:159–171. https://doi.org/10.2190/et.43.2.d

7. Fatt QK, Ramadas A. The usefulness and challenges of Big Data in healthcare. *J Health Commun*. 2018;**3(2)**. https://doi.org/10.4172/2472-1654.100131

8. Abdullah JM, inventors; Universiti Sains Malaysia, assignee. Big Data project on retrospective and prospective data for data hospital (pilot neuro). Intellectual Property Corporation of Malaysia, LY2018002680; 2018 July 10.

9. Abdullah JM. Neurosciences at Universiti Sains Malaysia represent Malaysia to support global precision brain health. *Malays J Med Sci*. 2019;**26(3)**:1–23. https://doi.org/10.21315/mjms2019.26.3.1

10. Global Brain Consortium. About us [Internet]. 2020 [Retrieved 2020 March 2]. Available at: https://globalbrainconsortium.org/

11. Daimiwal N, Sundhararajan M, Shriram R. Applications of fMRI for brain mapping. *International Journal of Computer Science and Information Security*. 2012;**10(11)**. Available at: https://www.researchgate.net/publication/234017771_Applications_of_fMRI_for_Brain_Mapping

12. Woo-Suk T, Shin-Hyuk K, Byung-Joo H, Byung-Jo K, Sung-Bom P. Brain mapping using neuroimaging. *Appl Microsc*. 2016;**46**:179–183. https://doi.org/10.9729/AM.2016.46.4.179

13. Pandarinathan G, Mishra S, Nedumaran A, Padmanabhan P, Gulyás B. The potential of cognitive neuroimaging: a way forward to the mind-machine interface. *J Imaging*. 2018;**4(5)**:70. https://doi.org/10.3390/jimaging4050070

14. Simon M. The four biggest challenges in brain simulation. *Nature.* 2019;**571(7766)**:S9. https://doi.org/10.1038/d41586-019-02209-z

15. Sumari P, Idris Z, Abdullah JM. We must invest in applied knowledge of computational neurosciences and neuroinformatics as an important future in Malaysia: the Malaysian brain mapping project. *Malays J Med Sci.* 2017;**24(1)**:1–9. https://doi.org/10.21315/mjms2017.24.1.1

16. Sapien Labs. Challenges of methodological variability in EEG [Internet]. Washington DC (USA); 2020 [Retrieved 2020 March 8]. Available at: https://sapienlabs.org/challenges-of-methodological-variability-in-eeg/

17. Newson J, Thiagarajan T. EEG frequency bands in psychiatric disorders: a review of resting state studies. *Front Hum Neurosci.* 2019;**12(521)**: 1–24. https://doi.org/10.3389/fnhum.2018.00521

18. Keil A, Debener S, Gratton G, Junghöfer M, Kappenman E, Luck S, et al. Committee report: publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. *Psychophysiology.* 2013:**51(1)**:1–21. https://doi.org/10.1111/psyp.12147

19. Parameshwaran D, Thiagarajan TC. Complexity of EEG reflects socioeconomic context and geofootprint. Review. 2016:1–18. https://doi.org/10.1101/125872