

Application of Big Data and Artificial Intelligence technologies to dementia prevention research: an opportunity for low-and-middle-income countries



Samuel O Danso¹, Graciela Muniz-Terrera¹, Saturnino Luz², Craig Ritchie¹; on behalf of the Global Dementia Prevention Program (GloDePP)

¹ Edinburgh Dementia Prevention, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, Scotland, UK

² Usher Institute, Edinburgh Medical School, Molecular, Genetic and Population Health Sciences, University of Edinburgh, Edinburgh, Scotland, UK

It is estimated that about 66% of the number of people with dementia globally live in regions within the low-and-middle-income countries (LMIC). The estimates further suggest that about 1 million additional cases are recorded annually in these LMICs, with approximately 60-70% of the dementia cases being due to Alzheimer disease, AD [1]. Despite this high incidence of AD, treatment options are currently limited. This is notwithstanding the fact that numerous clinical trials have taken place, are still on going or planned [2]. However, the rate of failure of clinical trials and predictions based on available results from randomised control trials suggest that an immediate breakthrough to obtaining treatment drugs for AD is unlikely [3].

While the symptoms of dementia appear to be well known, including decline in activities of daily living and social functioning, the exact cause remains unknown [4]. However, there is strong evidence based on several epidemiological studies to suggest that complex interactions exist between exposures such as adverse environmental factors and lifestyle choices and how these contribute to the risk and timing of dementia onset [5,6]. A comprehensive account of the various risk factors that interact to influence the risk profile and the timing of dementia onset is provided in the recently published Lancet Commission report [4]. The report furthermore suggests that the effect of these risk factors in terms of their contribu-

tion to dementia are beginning to be understood through complex modelling. These risk factors can be categorised into modifiable (such as lifestyle) and non-modifiable (such as genetics). The ability to alter the modifiable risk factors has been demonstrated theoretically to have influence on the onset of dementia. For example, it may be possible to reduce dementia prevalence by 50% if the onset was to be delayed by 5 years based on population-attributable risk estimation [7].

In the absence of strong epidemiological data from LMIC, evidence from numerous epidemiological studies carried out in high-income-countries (HIC) point to the fact that prevention is vitally important to reducing

Establishing norms for low- and middle-income countries' populations for these risk factors and developing monitoring mechanisms capable of providing early warning signs is an avenue to the development of strategies that can focus on early interventions.



Photo: LMICs leapfrogging with collaborative support and lessons learned from HICs to tackle dementia (photo taken from wpclipart collections: used under the terms and conditions).

the dementia burden. These findings may be extended to LMIC and this article discusses these opportunities and how they can be applied.

MODIFIABLE RISK FACTORS

It has now been established that the onset of dementia begins much earlier in life than had previously been assumed [5]. This is further supported by evidence from clinicopathologic studies, which suggest that AD lesions develop much earlier in specific regions of the brain which meet neuropathological AD diagnosis criteria even in the absence of dementia [8]. Meanwhile, numerous studies have also reported the effect of modifiable risk factors on the structural integrity of the brain. Examples include a 20-year longitudinal study, which found that cardiovascular risk factors such as hypertension, BMI from midlife to late life also increased the risk of severe white matter lesion in the brain [9]. A recent study also found a correlation between hypertension and reduction in brain reserve [10], and therefore adopting a healthy lifestyle may reduce the chances of hypertension which in turn also delay the onset of AD. Similar trends have been observed in nutrition studies [11]. Lifestyle factors such as physical activity through longitudinal observational studies have similarly been demonstrated to have improved cognitive reserves and reduction in dementia risk [12]. Furthermore, studies on lifestyle have also found smoking and excessive alcohol

intake to have significant association with cognition and dementia [13], with findings from education also showing some mixed effect on cognition and cognitive preservations [14]. Establishing norms for LMIC populations for these risk factors and developing monitoring mechanisms capable of providing early warning signs is an avenue to the development of strategies that can focus on early interventions.

DEMENTIA EARLY INTERVENTION THROUGH BIG DATA AND AI IN LMIC

HIC continue to explore advanced techniques in neuroimaging as well as neuropsychological and other data sources to effectively monitor brain health to detect early onset of dementia [15,16]. Recent advances in Big Data and AI technologies coupled that with the ever-increasing speed in data generation has seen exponential growth in research and development of these technologies within the context of dementia prevention. Researchers in HIC have already taken advantage of Big Data technologies and have developed data driven approaches to dementia prevention initiatives across various HICs and are able to process and manage these data with high throughput [17]. Current examples of such initiative include the ongoing European Prevention of Alzheimer Disease, EPAD [18] and PREVENT research programme [5]. These projects aim to generate not only large and high quality, but also phenotypically deep, data sets and are employing state-of-the-art Big Data technologies to process and make these data available through secured analytics environment for hypothesis testing.

Similar trends have also been observed in the analytics space where AI including sophisticated Machine Learning (ML) algorithms and frameworks such as deep learning continue to be developed and improved. A systematic review by Pellegrini et al. [19] found over 110 publications on various initiatives where ML approaches have been employed to develop prediction models for cognitive impairment and AD using neuroimaging data. These research efforts are being extended to other data sources and domains such as linguistic analysis of text messages; speech analysis [20,21]; and also through the human eyes using retinal imaging [22] with promising results.

While ML models based on neuroimaging data may not be practical in LMIC due to the huge cost associated with them, having alternative low cost approaches to acquiring modifiable risk factors data and analysing these data sources to help monitor populations brain health for the purposes of screening for dementia risks would be a useful epidemiological tool. For example, wearable devices and smartphones are now able to acquire real time data on daily activities, conversations through text messages and speech. The proliferation of the internet and smartphone usage in LMIC presents great opportunities for collecting these types of data using these devices. Big Data technologies allow efficient integration of data that come from a variety of sources and in different types and formats scale [23]. Harmonising these sources of data provides an excellent opportunity to develop ML methods and prediction models. AI and machine

learning technologies are also revolutionising the approaches to the analysis of large volumes of data in real time, with potential benefits to dementia research, including the ability to learn patterns from large data sets where use of traditional statistical methods may not be possible [24]. Additionally, the efficiencies and consistency in the use of AI-based methods may result in reduced cost and less human errors in decision-making.

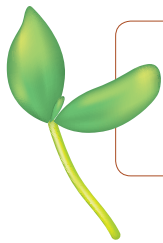
REGULATORY ISSUES

It is an undeniable fact that Big Data and AI have potential ethical challenges such as data privacy and respect for human rights. This has led to a shift in focus and a new branch of ethics known as data ethics, which involves the study of moral issues around data, algorithms and practices [25]. Regardless of the place of application, these ethical issues remain universal and applicable in LMIC. Nevertheless, regulatory frameworks such as the recently launched General Data Protection Regulations in May 2018 to regulate the use of personal data across Europe [26] and practises employed in other HIC such as the consent approach where study participants consent to researchers keeping their personal data centrally for the purpose of data linkage but these are not transferred to the data users. Another approach is the de-identification method where personal identifiable data are removed at the point of data collection before further processing and transmission as described in [23]. These approaches can serve as good examples to follow in LMIC.

Low- and middle-income countries can employ a leapfrog approach to adapt in order to benefit from these Big Data and AI technologies, which have already been tried and tested in high-income countries.

CONCLUSION

This article discussed opportunities for development of early intervention strategies to address the burden of dementia in LMIC through the application of Big Data and AI. The potential efficiency and cost-effectiveness of this approach suggests that window of opportunity now exists for researchers in LMIC to collaborate with colleagues in HIC in this area. LMIC can employ a leapfrog approach to adapt in order to benefit from these Big Data and AI technologies, which have already been tried and tested in HIC. With this approach, the emerging success stories in HIC can also be replicated in LMIC.



Funding: None.

Authorship contributions: SD drafted the manuscript with inputs and review from co-authors.

Conflict interest: The authors completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available upon request from the corresponding author), and declare no conflicts of interest.

REFERENCES

- 1 Prince M, Acosta D, Albanese E, Arizaga R, Ferri CP, Guerra M, et al. Ageing and dementia in low and middle income countries-Using research to engage with public and policy makers. *Int Rev Psychiatry*. 2008;20:332-43. [Medline:18925482](#) [doi:10.1080/09540260802094712](#)
- 2 Andrieu S, Coley N, Lovestone S, Aisen PS, Vellas B. Prevention of sporadic Alzheimer's disease: lessons learned from clinical trials and future directions. *Lancet Neurol*. 2015;14:926-44. [Medline:26213339](#) [doi:10.1016/S1474-4422\(15\)00153-2](#)
- 3 Schneider LS, Mangialasche F, Andreasen N, Feldman H, Giacobini E, Jones R, et al. Clinical trials and late-stage drug development for Alzheimer's disease: an appraisal from 1984 to 2014. *J Intern Med*. 2014;275:251-83. [Medline:24605808](#) [doi:10.1111/joim.12191](#)
- 4 Livingston G, Sommerlad A, Orgeta V, Costafreda SG, Huntley J, Ames D, et al. Dementia prevention, intervention, and care. *Lancet*. 2017;390:2673-4. [Medline:28735855](#) [doi:10.1016/S0140-6736\(17\)31363-6](#)
- 5 Ritchie CW, Ritchie K. The PREVENT study: a prospective cohort study to identify mid-life biomarkers of late-onset Alzheimer's disease. *BMJ Open*. 2012;2:e001893. [Medline:23166135](#) [doi:10.1136/bmjopen-2012-001893](#)
- 6 Imtiaz B, Tolppanen AM, Kivipelto M, Soininen H. Future directions in Alzheimer's disease from risk factors to prevention. *Biochem Pharmacol*. 2014;88:661-70. [Medline:24418410](#) [doi:10.1016/j.bcp.2014.01.003](#)
- 7 Norton S, Matthews FE, Barnes DE, Yaffe K, Brayne C. Potential for primary prevention of Alzheimer's disease: an analysis of population-based data 2014. *Lancet Neurol*. 2014;13:788. [Medline:25030513](#) [doi:10.1016/S1474-4422\(14\)70136-X](#)

- 8 Vuorinen M, Solomon A, Rovio S, Nieminen L, Käreholt I, Tuomilehto J, et al. Changes in vascular risk factors from mid-life to late life and white matter lesions: a 20-year follow-up study. *Dement Geriatr Cogn Disord*. 2011;31:119-25. [Medline:21273771](#) [doi:10.1159/000323810](#)
- 9 Killiany RJ, Hyman BT, Gomez-Isla T, Moss MB, Kikinis R, Jolesz F, et al. MRI measures of entorhinal cortex vs hippocampus in preclinical AD. *Neurology*. 2002;58:1188-96. [Medline:11971085](#) [doi:10.1212/WNL.58.8.1188](#)
- 10 Jeon SY, Byun MS, Yi D, Lee JH, Choe YM, Ko K, et al. Influence of hypertension on brain amyloid deposition and Alzheimer's disease signature neurodegeneration. *Neurobiol Aging*. 2019;75:62-70. [Medline:30553154](#) [doi:10.1016/j.neurobiolaging.2018.11.001](#)
- 11 Power R, Prado-Cabrero A, Mulcahy A, Howard A, Nolan JM. The role of nutrition for the aging population: Implications for cognition and Alzheimer's Disease. *Annu Rev Food Sci Technol*. 2019;10:619-39. [Medline:30908950](#) [doi:10.1146/annurev-food-030216-030125](#)
- 12 Blondell SJ, Hammersley-Mather R, Veerman JL. Does physical activity prevent cognitive decline and dementia?: A systematic review and meta-analysis of longitudinal studies. *BMC Public Health*. 2014;14:510. [Medline:24885250](#) [doi:10.1186/1471-2458-14-510](#)
- 13 Ngandu T, Helkala EL, Soininen H, Winblad B, Tuomilehto J, Nissinen A, et al. Alcohol drinking and cognitive functions: Findings from the Cardiovascular Risk Factors Aging and Dementia (CAIDE) study. *Dement Geriatr Cogn Disord*. 2007;23:140-9. [Medline:17170526](#) [doi:10.1159/000097995](#)
- 14 Wilson RS, Yu L, Lamar M, Schneider JA, Boyle PA, Bennett DA. Education and cognitive reserve in old age. *Neurology*. 2019;92:e1041-50. [Medline:30728309](#) [doi:10.1212/WNL.0000000000007036](#)
- 15 Ritchie K, Carrière I, Howett D, Su L, Hornberger M, O'Brien JT, et al. Allocentric and egocentric spatial processing in middle-aged adults at high risk of late-onset Alzheimer's Disease: The PREVENT Dementia Study. *J Alzheimers Dis*. 2018;65:885-96. [Medline:30103333](#) [doi:10.3233/JAD-180432](#)
- 16 Mortamais M, Ash JA, Harrison J, Kaye J, Kramer J, Randolph C, et al. Detecting cognitive changes in preclinical Alzheimer's disease: A review of its feasibility. *Alzheimers Dement*. 2017;13:468-92. [Medline:27702618](#) [doi:10.1016/j.jalz.2016.06.2365](#)
- 17 Doubal FN, Ali M, Batty DG, Charidimou A, Eriksdotter M, Hofmann-Apitius M, et al. Big data and data repurposing - using existing data to answer new questions in vascular dementia research. *BMC Neurol*. 2017;17:72. [Medline:28412946](#) [doi:10.1186/s12883-017-0841-2](#)
- 18 Solomon A, Kivipelto M, Molinuevo JL, Tom B, Ritchie CW; EPAD Consortium. European prevention of Alzheimer's Dementia Longitudinal Cohort Study (EPAD LCS): study protocol. *BMJ Open*. 2019;8:e021017. [Medline:30782589](#) [doi:10.1136/bmjopen-2017-021017](#)
- 19 Pellegrini E, Ballerini L, Hernandez MDCV, Chappell FM, González-Castro V, Anblagan DS, et al. Machine learning of neuroimaging for assisted diagnosis of cognitive impairment and dementia: A systematic review. *Alzheimers Dement (Amst)*. 2018;10:519-35. [Medline:30364671](#) [doi:10.1016/j.dadm.2018.07.004](#)
- 20 Orimaye SO, Wong JSM, Golden KJ, Wong CP, Soyiri NI. Predicting probable Alzheimer's disease using linguistic deficits and biomarkers. *BMC Bioinformatics*. 2017;18:34. [Medline:28088191](#) [doi:10.1186/s12859-016-1456-0](#)
- 21 de la Fuente Garcia S, Ritchie CW, Luz S. Protocol for a conversation-based analysis study: PREVENT-ED investigates dialogue features that may help predict dementia onset in later life. *BMJ Open*. 2019;9:e026254. [Medline:30918035](#) [doi:10.1136/bmjopen-2018-026254](#)
- 22 Csincsik L, MacGillivray TJ, Flynn E, Pellegrini E, Papanastasiou G, Barzegar-Befroei N, et al. Peripheral retinal imaging biomarkers for Alzheimer's Disease: A pilot study. *Ophthalmic Res*. 2018;59:182-92. [Medline:29621759](#) [doi:10.1159/000487053](#)
- 23 Deetjen U, Meyer ET, Schroeder R. Big data for Advancing Dementia research: an evaluation of data sharing practices in research on age-related neurodegenerative diseases, 2015. OECD Digital Economy Papers, No. 246, OECD Publishing. Available: <https://www.oecd-ilibrary.org/docserver/5js4sdbdf7jk-en.pdf?expires=1575056588&rid=id&accname=guest&checksum=91214D7FDBC5665BA4B8152516600565>. Accessed: 29 November 2019.
- 24 Breiman L. Statistical modeling: The two cultures. *Stat Sci*. 2001;16:199-231. [doi:10.1214/ss/1009213726](#)
- 25 Floridi L, Taddeo M. What is data ethics? *Phil. Trans. R. Soc. A*. 2016;374:20160360. [Medline:28336805](#) [doi:10.1098/rsta.2016.0360](#)
- 26 Li H, Yu L, He W. The impact of GDPR on global technology development. *J Global Inform Tech Manag*. 2019;22:1-6. [doi:10.1080/1097198X.2019.1569186](#)

Correspondence to:

Samuel O Danso, PhD
 Edinburgh Dementia Prevention
 Centre for Clinical Brain Sciences
 University of Edinburgh
 9A Bioquarter
 Little France
 Edinburgh, EH16 4 UX
 Scotland, UK
samuel.danso@ed.ac.uk