DigitCult | Scientific Journal on Digital Cultures

Deep Machine Learning Application to the Detection of Preclinical Neurodegenerative Diseases of Aging

Mathew J Summers
Sunshine Coast Mind & Neuroscience
Thompson Institute
University of the Sunshine Coast
Birtinya, Queensland, Australia

Tamas Madl
Austrian Research Institute for Artificial Intelligence (OFAI)
Vienna, Austria

Alessandro E. Vercelli
Neuroscience Institute Cavalieri Ottolenghi
Department of Neuroscience
University of Turin, Turin, Italy

Georg Aumayr
Johanniter Austria Research and Innovation Centre
Österreich, Vienna, Austria

Doris M. Bleier
Johanniter Austria Research and Innovation Centre
Österreich, Vienna, Austria

Ludovico Ciferri
International University of Japan
Niigata-ken, Japan

Abstract

Artificial intelligence (AI) deep learning protocols offer solutions to complex data processing and analysis. Increasingly these solutions are being applied in the healthcare field, most commonly in processing complex medical imaging data used for diagnosis. Current models apply AI to screening populations of patients for markers of disease and report detection accuracy rates exceeding those of human data screening. In this paper, we explore an alternate model for AI deployment, that of monitoring and analysing an individual's level of function over time. In adopting this approach, we propose that AI may provide highly accurate and reliable detection of preclinical disease states associated with aging-related neurodegenerative diseases. One of the key challenges facing clinical detection of preclinical phases of diseases such as dementia is the high degree of inter-individual variability in aging-related changes to cognitive function. AI based monitoring of an individual over time offers the potential for the early detection of change in function for the individual, rather than relying on comparing the individual's performance to population norms. We explore an approach to developing AI platforms for individual monitoring and preclinical disease detection and examine the potential benefits to the stakeholders in this technological development.

Keywords

artificial intelligence; deep learning; dementia; mild cognitive impairment; ageing; aging; neurodegenerative disease; diagnosis; preclinical; prodrome

Published 11 December 2017

Correspondence should be addressed to A/Prof Mathew Summers, Sunshine Coast Mind & Neuroscience – Thompson Institute, University of the Sunshine Coast, Locked Bag 4 (ML59), Maroochydore DC, Queensland, Australia, 4558. Email: msummers@usc.edu.au

DigitCult, Scientific Journal on Digital Cultures is an academic journal of international scope, peer-reviewed and open access, aiming to value international research and to present current debate on digital culture, technological innovation and social change. ISSN: 2531-5994. URL: http://www.digitcult.it

 DOI:10.4399/97888255088952

Artificial Intelligence in Healthcare – From the Present to the Future

Artificial intelligence (AI) is an umbrella term that refers to computers, robots, and software that mimic aspects of human intelligence. AI is of increasing importance in the healthcare sector. A key driver in the increase in AI-based data-driven healthcare is the need for the global healthcare industry to reduce costs and more efficiently manage resources while improving patient care. In addition, the increasing prevalence of chronic diseases, aging populations, changing consumer expectations about how they want to purchase and receive care, and increasing access to social media and mobile technologies are transforming the way healthcare is obtained and delivered. In fact, there are already several examples of AI applications developed for healthcare, with the number of startups entering the healthcare AI space increasing in recent years, with over 50 companies raising their first equity rounds since January 2015. Deals to healthcare-focused AI startups increased from less than 20 in 2012 to nearly 70 in 2016. In 2016 two new unicorns emerged into the space: China-based iCarbonX and oncology-focused Flatitron Health (CB Insights 2017).

One of the many advantages of AI is that decisions made by AI are evidence-based and devoid of emotional influence and cognitive biases. This differentiates AI decision-making processes from human decision making processes which incorporate an inherent intuitive/emotive component in formulating a response. Instead, for the AI system, intuitive-thinking is a product of the analysis of past behaviors of its user and large data sets. In order to implement AI into the healthcare industry there is a need for healthcare organizations to become data-driven, treating data as strategic assets and implementing processes and systems to inform decision-making procedures and drive actionable results.

Data gathering for machine learning and deep learning capabilities have immense potential to improve diagnostics, care pathway creation and reproducibility in surgical procedures to ultimately achieve better clinical outcomes. For example, deep learning can involve the use of wearable technology targeted to specific conditions, such as remote monitoring of cardiac function of an individual, using individually-tailored algorithms derived from the individual's biometric and patient data. The utilization of machine learning principles can advance the analysis of the unstructured data delivered from these medical-grade wearable devices to clinically relevant diagnostic information, with mathematical algorithms trained to detect anomalies in this data. Machine learning-based decision support systems could then interpret the meaning of anomalies, in a way similar to an expert human physician, resulting in a considerable saving in physician time that would otherwise be expended in processing unstructured data. As such, AI based platforms undertaking high level analysis of unstructured data enables the healthcare professions to focus on treatment delivery to the most critical patients and streamline the care process. According to a new research report (MarketsandMarkets Analysis 2017), the worldwide market is expected to grow from USD 667.1 Million in 2016 to USD 7,988.8 Million by 2022, at a CAGR of 52.68% during the forecast period. Another report, slightly more conservative in its modelling, is nonetheless indicating that the growth in the AI health market is expected to reach USD 6,600 Million by 2021, which represents a 40% compound annual growth rate (Collier, Fu, and Yin 2017).

Current approaches to integrating AI platforms into disease modelling mainly undertake a traditional disease focused approach. Instead, AI could be a potential solution to the medical scientific challenge of detecting preclinical neurodegenerative disease by deep learning protocols to develop population based screening of disease states. Automatic detection of preclinical diseases from a multitude of sensors presents several challenges to traditional machine learning, yet would be in line with the future of AI in medical technology: that of individuated monitoring and intervention.

Neurodegenerative Disease: The Current Challenge in the Diagnosis and Detection of Preclinical Phases

Neurodegenerative diseases encompass a cluster of hereditary or sporadic conditions characterised by degeneration of central and/or peripheral neural tissue. There are over 600 known human neurodegenerative diseases of different aetiologies and trajectories, with some being terminal conditions with few treatment options. The majority of cases of neurodegenerative disease are a result of dementia, of which there are at least 50 different known variants. The most prevalent neurodegenerative disease is Alzheimer's Dementia (AD) (Alzheimer's Association 2017), with other relatively common neurodegenerative disorders including Vascular Dementia (VaD), and Parkinson's Disease (PD). An estimated 10% of adults aged 65 years and older have AD, with the prevalence of AD increasing with age (17% of 75-84 yr olds; 32% o 85+ yr olds) (Alzheimer's Association 2017).

A major challenge in the field of neurodegenerative diseases is early and timely diagnosis of the disease. For the majority of neurodegenerative diseases the underlying causal agent for the disease remains unknown. Consequently, there are no definitive biologic tests to detect the presence or absence of the disease *in vivo*. The *in vivo* (*ante-mortem*) diagnosis of aging-related neurodegenerative diseases (dementias, etc.) is made on the basis of the presence of hallmark clinical features, encompassing loss of functional capacity alongside recognised clusters of cognitive, physical, social, and/or mood disturbances (McKhann et al. 2011). In the case of Alzheimer's dementia, loss of capacity to maintain independent living (functional impairment) with progressive deterioration of memory (new learning as well as access to old information), language processing (typically fluent aphasic conditions), spatial orientation, and executive functions (attention, concentration, decision-making, planning) are characteristic clinical symptoms of AD (McKhann et al. 2011, Alzheimer's Association 2017).

These clinically diagnosed cases are associated with post-mortem neuropathology. In the case of AD, the presence of neurofibrillary tangles and senile plaques are considered pathological hallmark features (Braak and Del Tredici 2012). However, plaques and tangles are not unique to AD and are evident in normal aging as well as in other age-related diseases – the hallmark feature for AD is the distribution and density of the plaques and tangles. Further, not all cases with *post-mortem* pathology consistent with AD display *ante-mortem* clinical symptoms of AD (Mufson et al. 2016). That is to say, plaques and tangles reflect the consequence of a disease agent and are not the disease agent (Drachman 2013, Hardy and De Strooper 2017). The past 20 years has seen an increase in the use and development of imaging techniques and biomarker assessment of potential *ante-mortem* disease markers for AD. These include MRI imaging (e.g., hippocampal volumetric analysis), amyloid binding in PET imaging, functional activation changes on fMRI, beta-amyloid load in CSF, etc. To date, none of these techniques has displayed sufficient sensitivity and specificity to reach the threshold required for early diagnosis of AD (McKhann et al. 2011). Therefore, the clinical diagnosis of AD remains based on clinical symptomatology, with biomarkers and imaging providing exclusionary/confirmatory diagnostic information.

With this lack of knowledge regarding the causal agent(s) for AD, an effective treatment for the disease AD will remain elusive. Pharmacological treatments currently being developed and tested predominantly target the biochemical pathways associated with accumulation of the properties of tangles and plaques (i.e., beta-amyloid deposition), without targeting the trigger for this deposition. At best, these approaches may offer treatment of the symptoms of AD without treating the disease itself. However, recent clinical trials of initially promising drugs to reduce beta-amyloid load have returned disappointing or negative treatment effects resulting in the cessation of the majority of clinical trials (Abbott and Dolgin 2016, Hardy and De Strooper 2017).

In the past 15 years there has been a resurgence in the attempt to identify preclinical markers of diseases such as AD. The most recent attempt includes the notion of mild cognitive impairment (MCI) initially identified by the Mayo Clinic on the basis of a cohort study tracking adults into AD (Petersen et al. 1999). The Mayo Clinic study identified that prior the emergence of clinical symptoms of AD, subclinical decline in memory was observed in those that subsequently developed AD (Petersen et al. 1999). Similar findings have been found around the world, increasing interest in the detection of MCI as a preclinical phase of AD. More recent research has highlighted problems with the sensitivity and specificity of MCI, with longitudinal studies of populations identified as suffering from MCI reporting that up to 40-70% of MCI cases tracked longitudinally return to an unimpaired state, indicating that MCI lacks validity as a preclinical AD diagnosis (Summers and Saunders 2012, Klekociuk et al. 2014, Delano-Wood et al. 2009,

Edmonds et al. 2015, Edmonds et al. 2016). Increasingly, researchers are acknowledging that MCI represents a risk factor for AD and not a diagnostic marker for preclinical AD.

The example provided here of the challenges facing detection and diagnosis of AD and the impact that this poses with the detection and diagnosis of preclinical disease states such as MCI affects the majority of the age-related neurodegenerative diseases. The conundrum is self-evident; in the absence of a known causal agent for a disease, the capacity to diagnose the disease is dependent solely on the validity and accuracy of the clinical features of the disease and not the presence or absence of a disease causing agent. The signs and symptoms of a clinical disease state are dependent on the homogeneity of the population with the disease; variability in symptomatology or underlying causal agents decreases diagnostic accuracy of the clinical signs and symptoms. Therefore, the clinical diagnosis of AD is dependent on two key assumptions:

1. There is a universal common causal agent for the disease – if the syndrome (e.g. AD) has multiple causal agents, then it is likely that the clinical diagnosis of AD is not a disease specific diagnosis but captures different diseases with similar clinical symptoms;

2. That all persons with the disease display the same clinical signs and symptoms. There is increasing evidence of variability in symptom presentation and temporal changes in symptomatology in persons with a clinical diagnosis of AD.

Longitudinal studies of aging indicate considerable variability in the trajectories of functional change over time (see Figure 1) (Finkel, Ernsth-Bravell, and Pedersen 2015, Ylikoski et al. 1999). Typically, longitudinal studies address this variability by computing group averaged trajectories based on outcome, such as normal aging versus dementia as an outcome (see Figure 2). However, a group averaged trajectory masks the underlying variability, which when recognised highlights the discrepancy between the change over time of an individual's level of cognitive function compared to a group averaged trajectory of cognitive decline.

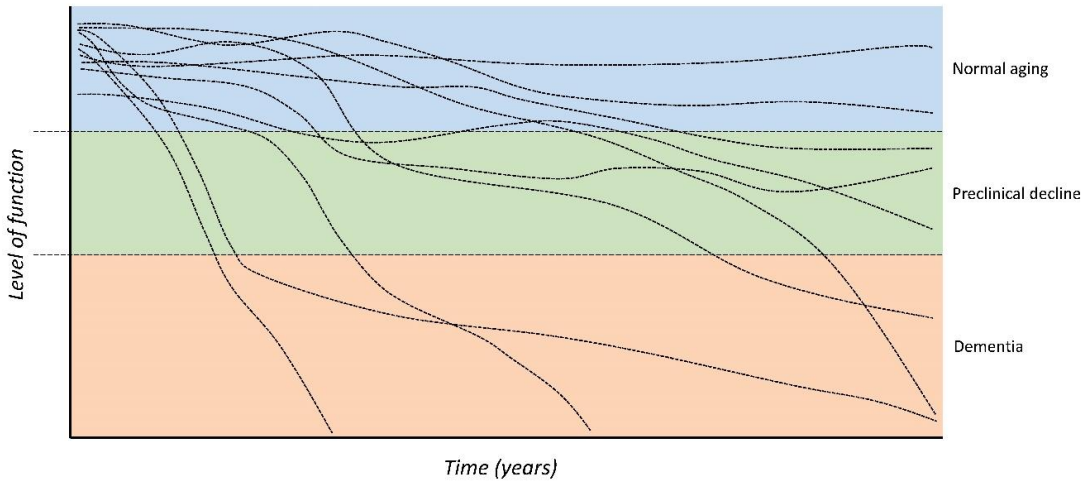

Figure 1. Individual trajectories of functional decline in aging over time.

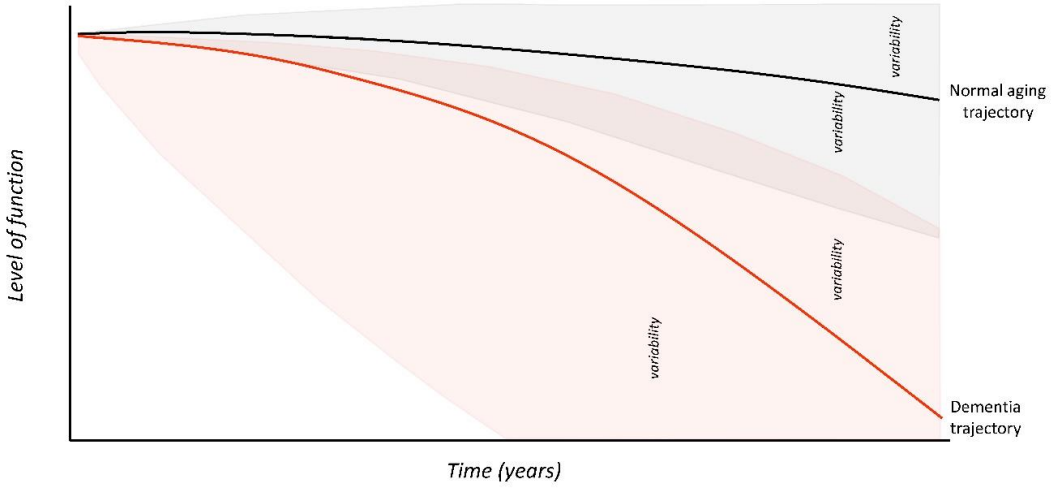

Figure 2. Outcome group averaged decline trajectories.

Thus, detection of preclinical disease states is extremely challenging where the disease agent remains unknown and may be multifactorial, where symptomatology of the clinical disease state is variable, and where the signs and symptoms of the clinical disease state are not necessarily manifest in the pre-clinical stages (otherwise clinical disease criteria would be met), and where diagnostic criteria lack adequate sensitivity and specificity due to multiple statistical challenges (Klekociuk, Saunders, and Summers 2016). Current preclinical models are derived from identifying common signs and symptoms in subpopulations who subsequently develop the clinical disease state of interest, but these represent potential risk factors and not diagnostic features (Klekociuk, Saunders, and Summers 2016).

AI as a Potential Solution to the Challenge of Detecting Preclinical Neurodegenerative Diseases

Current approaches to integrating AI platforms into disease modelling undertake a traditional disease focused approach (Woo et al. 2017). Typically, this approach:

1. is based on the premise of isolating new methods to identify disease states in population cohorts – essentially using AI to process data from large data sets from the population and detect statistical deviancy (in patterns of data) against a background of normal variance. The statistically deviant patterns are then correlated against an outcome metric (e.g. disease state) to demonstrate efficacy of using AI to screen the general population for diseases (Figure 3).

2. The goal of this approach is to identify individuals in the general population who display patterns of medical test results (etc) that indicate the presence of a disease state.

The ultimate aim of current approaches is to replace existing methods (i.e., expert human screening) with an AI screener with lower error rate for detection. AI development in this field requires very large data sets for deep machine learning approaches to be implemented. A major constraint on deep machine learning is variance in the patterns being examined. In the case of detection of abnormal pathology markers in a population set, the main source of variance is inter-individual variability, with additional sources of variance from variability in medical equipment (e.g. different in number and strength of the magnets used in MRI as well as the algorithms employed by the MRI software).

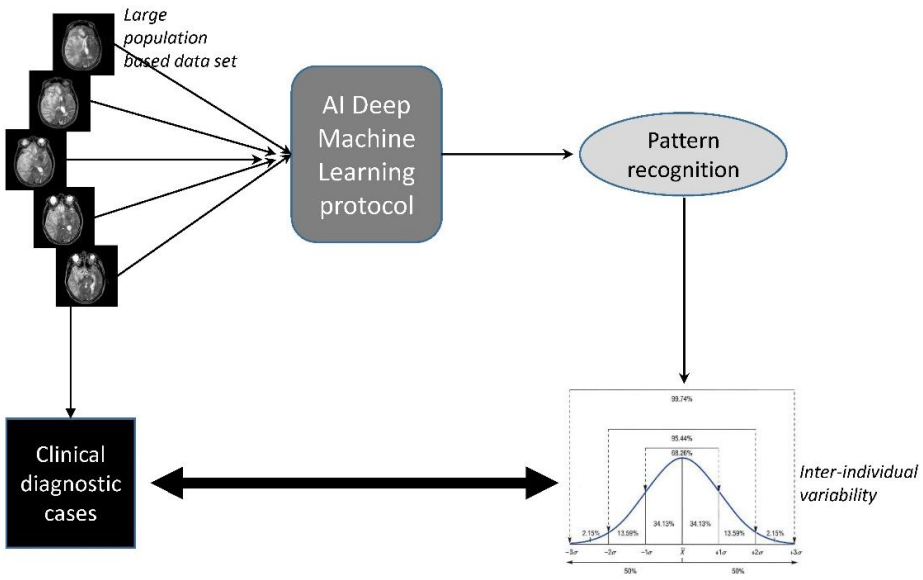

Figure 3. Artificial intelligence system to detect disease within a population.

However, the current approach of population based screening cannot lead to AI detection of preclinical disease states due to fundamental conceptual issues:

AI cannot diagnose a disease state independent of human bias. Clinical diagnoses are conceptual models developed through clinical practice that best describe the commonalities of signs and symptoms displayed by the majority of a subgroup of the wider population (the "afflicted"). It is important to note that there is variability in the signs and symptoms displayed by the afflicted, with the diagnostic criteria describing what the majority of the afflicted display. Therefore, AI pattern recognition can only identify a pattern of information that best matches the clinical diagnosis. The less precisely the clinical diagnosis matches the afflicted, the less accurately the AI will identify those with the disease state.

The constructs of preclinical disease states are purely conceptual and based on the premise that a precursor to frank disease must exist. For the majority of the "diagnosis" of a preclinical disease is based on milder forms of the signs and symptoms associated with the clinical disease state (e.g. mild cognitive impairment as a preclinical state of dementia). However, there is no good scientific evidence supporting the validity of preclinical diagnostic criteria.

The causal pathogenic agent of various neurodegenerative diseases (e.g. Parkinson's disease, AD, Frontotemporal Dementia, etc.) remain unknown. Hence, clinical diagnosis is not made on the basis of a direct measure of the causal agent, but rather a presence of a cluster of signs and symptoms that may or may not relate to a common causal agent. There is increasing debate as to whether clinical diagnostic states (such as Parkinson's dementia or AD) represent a homogenous disease; or rather represent a cluster of related diseases with different causal agents. That is, commonality of signs and symptoms does not necessarily indicate a common causation.

Can AI Solve the Challenge of Detecting Preclinical Neurodegenerative Disease States?

It is necessary to look beyond the use of AI deep learning protocols to develop population based screening of disease states and consider the future of AI in medical technology: that of individual monitoring and intervention. In this approach, a deep machine learning protocol is applied to a *large dataset derived from a single individual*. The volume of data is a direct product of continuous information collected from a single individual over an extended period of time. The volume of data required will be determined by the time taken for the AI deep machine learning protocol to learn and recognise patterns of data from a single person (Figure 4).

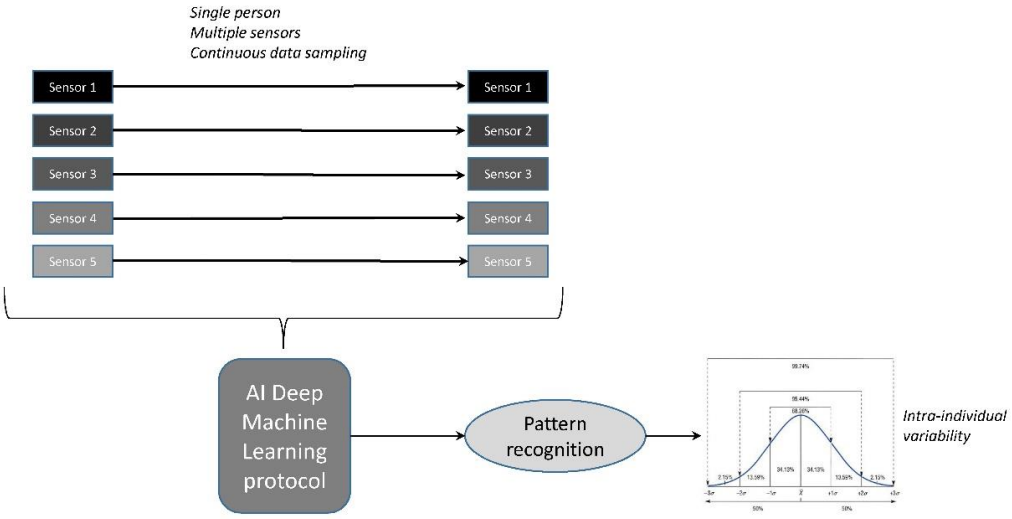

Figure 4. Artificial intelligence system to detect decline within an individual.

The Artificial Intelligence Challenge to Detecting Individual Change over Time

The modern AI methods often referred to as “deep learning” or “deep machine learning”, have been demonstrated to match or exceed human accuracy in classifying diseases such as: lung cancer (Yu et al. 2016) and brain gliomas (Ertosun and Rubin 2015) from tissue samples, breast lesions and pulmonary nodules from CT imaging (Cheng et al. 2016), Alzheimer’s disease from fMRI scans (Sarraf, Tofighi, and Alzheimer Disease Neuroimaging Initiative 2016) and MRI scans (Suk, Lee, and Shen 2017); and, skin cancers from photographs (Esteva et al. 2017). However, these methods are dependent on extremely large volumes of training data in order to reach a level of deep learning classification accuracy (see Figure 5). In the case of AI skin cancer detection, a total of 129,450 individual clinician-annotated photographs of skin cancers were required for AI training (Esteva et al. 2017).

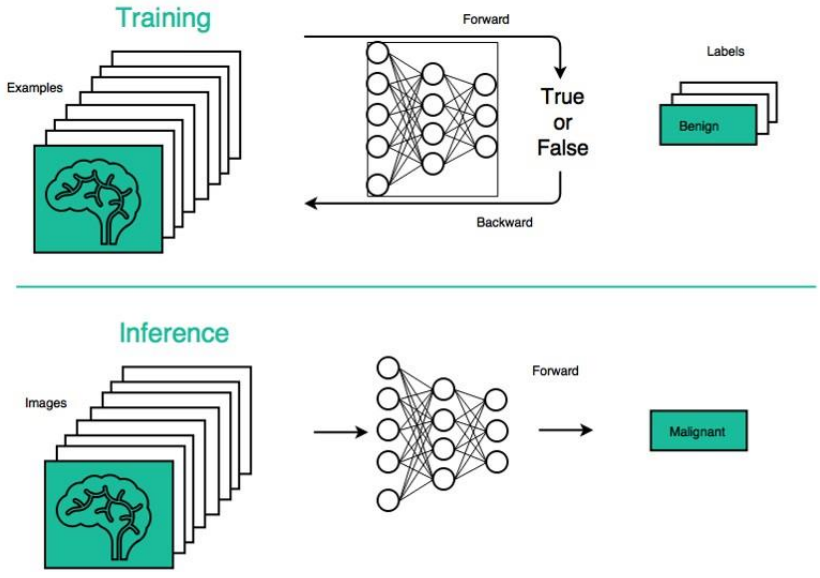

Figure 5. Deep learning training approach.

While the paradigm shift from massive labelled datasets for training supervised deep learning models to processing longitudinal data from individual patients represents a significant leap towards personalized screening and diagnosis of preclinical disease, there are several technical obstacles to be overcome for this idea to be feasible. Three major problems have to be solved in order to develop a system able to provide predictions in a reasonable time frame, without requiring the millions of labelled instances that traditional deep learning usually has access to:

1. The curse of dimensionality: multidimensional data can make learning difficult, particularly as the presence of irrelevant features frequently conceals the presence of anomalies;

2. The limited availability of data: modern deep neural networks function extremely well in settings with abundant labelled data points (famous examples such as Google's Inception network have learned from tens of millions of instances), but are frequently inferior to other techniques when datasets are small. Indeed, it can be proven that no single method is optimal for all possible settings and data dimensions (no free lunch theorem);

3. The difficulty of obtaining reliable labels: in the setting of longitudinal observation and screening for pre-clinical disease, reliable diagnosis by experts is both time consuming, expensive, and made more difficult by the lack of well-substantiated preclinical diagnostic criteria.

A strong model performing well in a real-world setting would need to autonomously pick out relevant features/dimensions for recognizing anomalies and trends; be capable of learning from limited data from a single individual; and rapidly reduce false positives and false negatives from a handful of expert opinions. These issues preclude taking existing off-the-shelf deep machine learning models and applying these models to preclinical disease screening, irrespective of how well they work in domains with abundant and accurately labelled datasets.

The field of anomaly detection has developed several methods for recognizing when a data stream or time series deviates from what is considered normal; from simple outlier detection methods based on a statistical model to the one class support vector machine (1-class SVM). However, all of these methods require a small number of reliable features, and break down in a high-dimensional setting where not all dimensions are relevant. The field of combining anomaly detection with dimensionality reduction methods to mitigate this problem is still in its infancy (Erfani et al. 2016).

The focus of machine learning has recently shifted from supervised classification to unsupervised (no labelled instances) and semi-supervised learning (partially labelled instances), and several recent breakthroughs have yielded practically applicable methods such as deep auto-encoders (unsupervised) or ladder networks (semi-supervised). However, these methods have been designed for "deep" (rich multivariate data source with narrow temporal distribution) and not "wide" (temporally distributed data sources) data; and they are not trivially applicable to lengthy time series from single patients.

Finally, physiological and biological modelling make it possible to incorporate large amounts of prior expert knowledge into computational models, thus significantly reducing the amount of experimental data required to fit a reliable model. However, the fusion of powerful machine learning methods with hand-crafted physiological models, which could combine the advantages of both, has been neglected in existing literature (Madl 2017).

In order to develop individuated AI monitoring, three approaches to deep-learning need to be built upon:

- Unsupervised representation learning methods, adapted to longitudinal multi-sensor data from individual patients to tackle the curse of dimensionality (together);

- Physiological / biological models with further constraints on the relevant dimensions by making use of expert feedback (Figure 6), in order to reduce the amount of data required for reliable learning, and combine prior scientific knowledge about a phenomenon with learning from raw data;

- Semi-supervised learning methods which classify a large number of phenomena based on very few expert labels, as well as information-theoretic feature selection to reduce the dimensionality with expert labels, in order to tackle the difficulty of obtaining a large number of reliable labels (Figure 7).

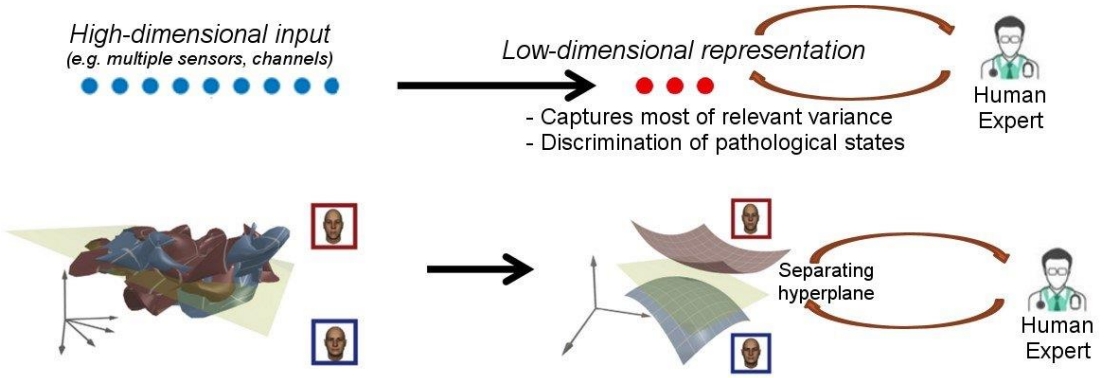

Figure 6. Combining data-driven machine learning with human expertise.

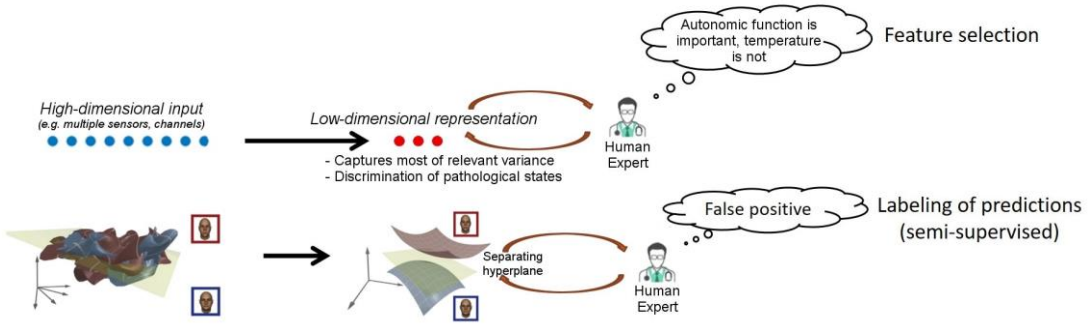

Figure 7. Semi-supervised machine learning.

Outline of a Possible Artificial Intelligence System for Learning Preclinical Decline

Early detection of preclinical decline is crucial for appropriate care and preventative measures, especially given the increasing incidence rates in an aging population. Although traditional tools providing insights into neurodegenerative diseases (such as brain imaging) are in general prohibitively expensive as population-wide screening measures, there exist cheap and ubiquitous sensors that can be used for this purpose do exist. However, no clear “biomarkers” in such consumer sensor data are known; markers of decline have to be autonomously acquired from individuals, in the longitudinal fashion outlined in the sections above. In practice, individuals will be constantly monitored along a number of dimensions (e.g., accelerometry, heart rate, SpO2, body temperature, voice / speech patterns, facial recognition of mood state, etc.), with consumer devices such as wearables or smart watches and smartphones equipped with the appropriate sensors.

In order to tackle the above-mentioned challenges associated with learning from such data, an unsupervised representation learning model (enhanced by physiological models) could combine multidimensional time series to a compressed feature space, in order to represent the vast majority of variance in the input data using a small number of relevant features, as well as reducing noise and irrelevant information. A good initial approach would be a deep unsupervised machine learning model such as a deep autoencoder. This will result in a concise representation space spanned by relevant dimensions acquired automatically from the input data, in such a way as to minimize reconstruction error.

Such an unsupervised representation learning approach is well suited to deal with the curse of dimensionality, and may allow the autonomous discovery of "biomarkers" that correlate with aspects of preclinical decline. For example, frailty frequently leads to impaired autonomic control (Varadhan et al. 2009), a correlation which could be detected autonomously from heart rate data, as the variability in the inter-beat intervals is very likely to be one of the important features discovered by an unsupervised representation learning model. To exploit such representations beyond simple correlation with clinical variables, critical events (such as falls, arrhythmias, or even cardiac arrest) could be recognized by applying anomaly detection algorithms, which are designed to identify deviations from the usual signal. Thus, an anomaly detection model (Figure 8) could operate on this compressed space, and identify detrimental trends or abnormal phenomena. Anomaly detection systems usually have a certain false positive rates, which have to be mitigated by a learning approach. To this end, identified abnormalities will initially be manually verified by clinicians, and the diagnostic information will feed back to the model, further constraining the feature space, and providing a semi-supervised learning signal (Figure 8). After a relatively short learning phase involving frequent supervision, such a system has the potential to acquire: (1) a compressed representation that can capture health states with a very small number of variables; (2) decision boundaries which accurately separates normal from potentially abnormal health states, and (3) physiological sub-models as part of the machine learning model which, apart from reducing the amount of training data required for reliable learning, may also provide crucial insights to clinicians in the future (Figure 9).

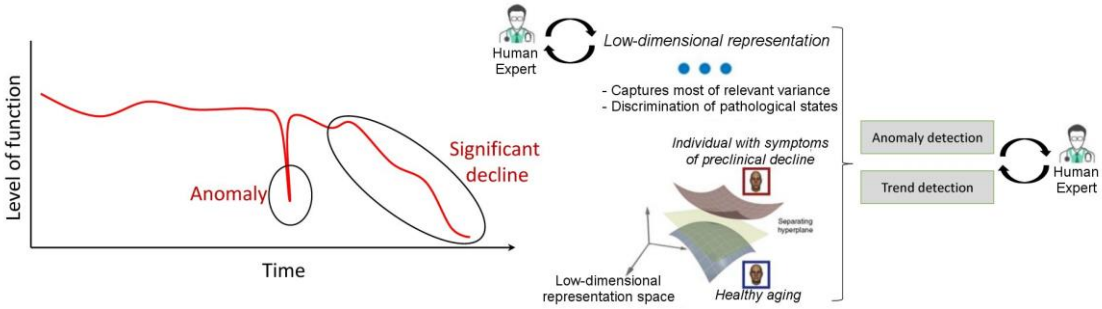

Figure 8. Detecting decline as a time series in a learned representation.

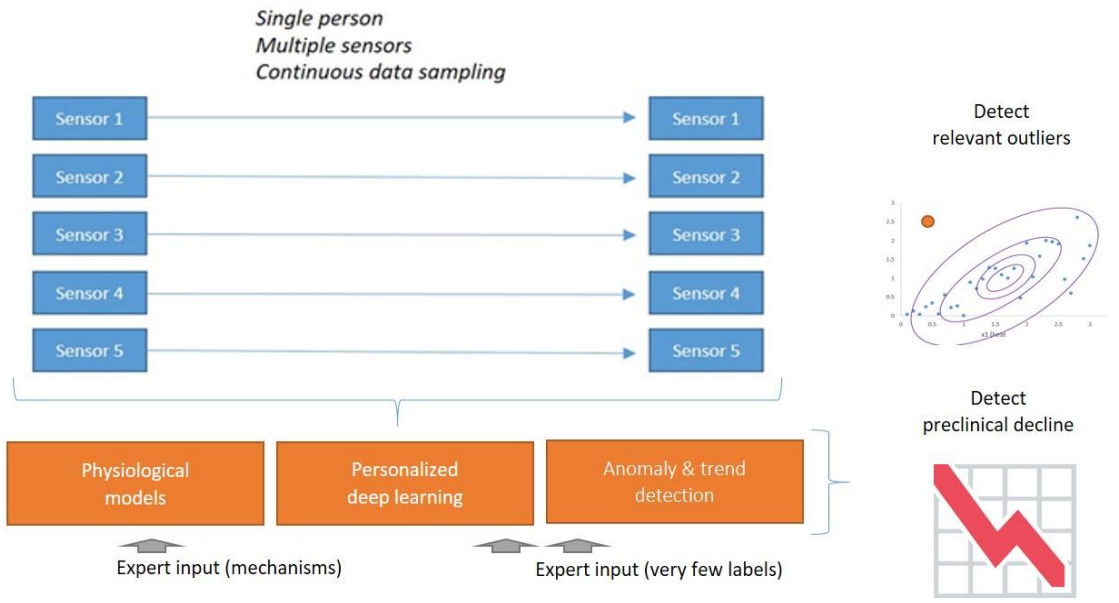

Figure 9. Potential AI architecture.

Previous research has demonstrated representation learning with little supervision enables AI platforms to learn and predict the structure of more than 250 human spatial memories (Madl, Franklin, Chen, Trappl, et al. 2016) and enabled a humanoid robot to learn spatial memories with an accuracy comparable to humans (Madl, Franklin, Chen, Montaldi, et al. 2016). AI models combining machine learning with physiological models have been used to detect stable coronary artery disease from inter-beat intervals (Madl 2016, 2017).

Impacts & Implications for the Stakeholder

AI has the potential to exert a major impact on healthcare provision for the stakeholder. The impacts are likely to be experienced on both diagnostic processes as well as the human workload for the provision of healthcare services. Examples of the potential for significant impact come from case studies of the use of AI in the areas of health documentation and adherence.

The need of support by ICT systems in healthcare is very clear. Worldwide, there is a population shift occurring, arising from improved healthcare in early life, the average of age of mortality is progressively shifting to later age groups over time. The pattern evident in changes to mortality in the Brussels region of Belgium (Figure 10) is replicated across most industrialized societies. However, it is important to note that an increase in lifespan does not necessarily indicate a commensurate increase in health levels within these populations. It is possible that with advancing age there is an increase in health-related problems that further add to the burden of care for the stakeholder.

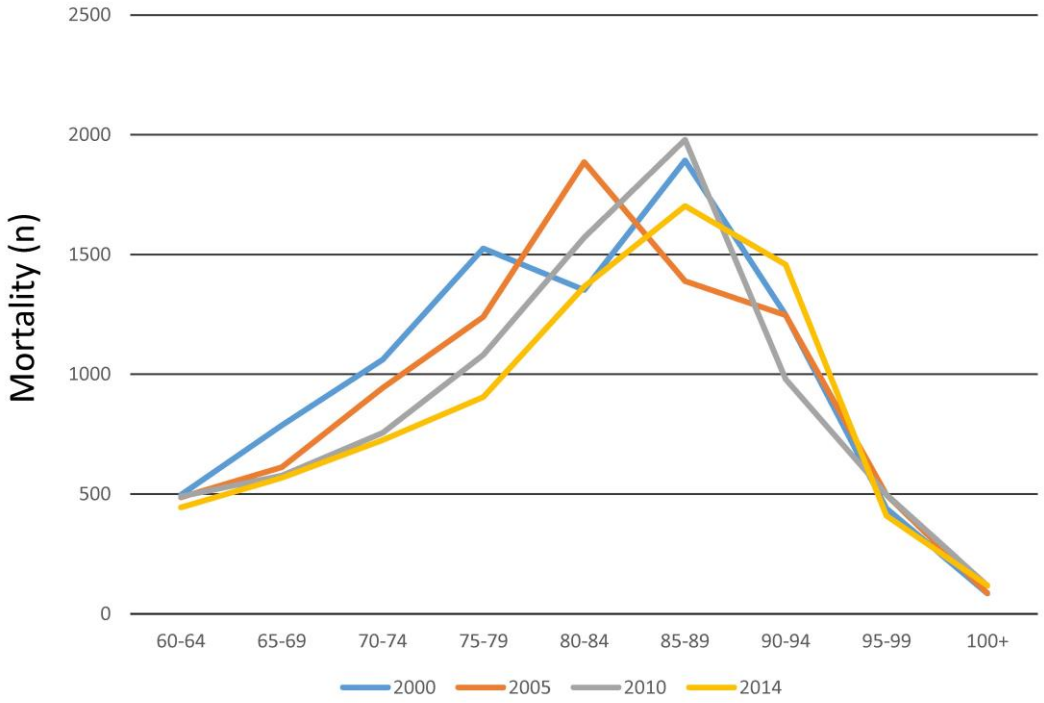

Figure 10. Mortality rate within age bands at different years for the region of Brussels, Belgium.

AI can directly contribute to the enhancement or improvement of several aspects of healthcare. For example, it can assist staff operating in high pressure environments, such as intense or prolonged working hours in situations requiring rapid critical decisions, by providing additional information for supporting routine decisions. Furthermore, an AI system can prove valuable when used for providing diagnostic options, as an AI platform can support decisions and identify changes in health status using evidence-based heuristics, where expensive diagnostic equipment is not available. Sensors can be deployed to provide data to the AI system to identify and monitor medication effects (adverse and desired) as well as monitor various behaviors. This can help in early risk detection, disease monitoring and therapy adjustment. AI can also help in increasing effectiveness in documentation and control of interventions. The benefits for end users include being able to monitor progress and enable basic data analysis. This data in turn can assist in raising awareness and combating physical or mental decline.

Metadata can be used to gain information on the time users spent interacting with the AI system (for example time spent walking, data added manually by an end user, time spent on exercises that aim to improve physical and mental wellbeing). Furthermore, it provides a long-term and ongoing monitoring tool compared to caregiver monitoring in which the patient may only be monitored sporadically or intermittently (once, twice, perhaps thrice a day). An AI system (for example on a mobile device with some additional sensors) can be customized to act (through a predefined prompt) and react (for example a predefined prompt when an individually defined heart rate is reached) to the needs of each end user and allows for monitoring throughout the day, regardless of whether the caregiver is currently present or not.

Documentation

At present, clinical documentation procedures account for approximately 15.8% of workload time for nurses (Korst et al. 2005). Of this 15.8%, a total of 10.6% is performed using paper-based recording systems. Over the coming years a reduction in paper-based records will occur as most

healthcare systems increasing deploy electronic-health record (EHR) systems, which has the potential to decrease the overall time required to perform and maintain healthcare documentation. The reduction in documentation time has the capacity to create additional staff time for clinical patient care which is an increasing demand on clinical staff workloads (e.g. because of inadequate staffing) (Neill 2011). Rather than using a centralized computer record system (Poissant et al. 2005) the use of personal devices enabling documentation at the patient bedside is seen as the most effective replacement of current paper-based record systems.

All data produced during the documentation process would be transmitted to the centralized record system and immediately available in the patient's EHR of the hospital. This makes the transfer of paper based documents to the electronic system obsolete and decreases the chance for forgetting important details in the documentation when doing it later.

Adherence

Ease-of-use and a high level of individualization will be key to ensuring that end-users remain active system users over a long period of time. The AI system must be able to assist patients (directly or through medical staff) with good physical and/or mental health as well as users with poor physical and/or mental health. Further, the system must be adaptive, for example by recognizing a trend as such and adapting the relevant thresholds accordingly. A main challenge for this kind of system will be the personalization of motivational cues to increase adherence. If this is successful, the costs in healthcare will decrease dramatically for individuals with chronic diseases. Being able to monitor progress and present collected data in a user-friendly way should be able to encourage adherence in a similar fashion by increasing belief in self-efficacy. Examples for this can be found in diabetes care (Littlefield et al. 1992) and treatment of hypertension (Breaux-Shropshire et al. 2012). Compliance can similarly be encouraged through motivation and thus needs to be communicated in a supportive and encouraging way.

Conclusion

There is a clear need for novel solutions to mitigate aging-related diseases, which are known to have steadily increasing healthcare costs as well as significant but modifiable losses to quality-adjusted life years. Artificial intelligence methods are beginning to make an impact on the healthcare sector, but have predominantly been used in settings characterized by clearly defined, distinct categories of health states (such as cancerous or non-cancerous tissue), as well as the availability of large amounts of data across many individuals.

In this paper, we have argued for a fundamentally different approach: using AI models applied to large datasets derived from single individuals, in order to detect preclinical decline. We have summarized some of the systemic and technical challenges associated with this task, and outlined approaches to tackle them based on recent AI research and techniques, and their potential impact and implications. Future research, not limited to mainstream deep learning methods, will be required to arrive at mature solutions.

References

Abbott, Alison, and Elie Dolgin. "Leading Alzheimer's theory survives drug failure." *Nature* 540 (2016): 15-16.

Alzheimer's Association. "2017 Alzheimer's disease facts and figures." *Alzheimer's & Dementia* 13.4 (2017): 325-373. doi: http://dx.doi.org/10.1016/j.jalz.2017.02.001

Braak, Heiko, and Kelly Del Tredici. 2012. "Alzheimer's disease: Pathogenesis and prevention." *Alzheimer's & Dementia* 8 (3):227-233.

Breaux-Shropshire, Tonya L., Kathleen C. Brown, Erica R. Pryor, and Elizabeth H. Maples. "Relationship of blood pressure self-monitoring, medication adherence, self-efficacy, stage of change, and blood pressure control among municipal workers with hypertension." *Workplace Health & Safety* 60.7 (2012): 303-311. doi: http://dx.doi.org/10.1177/216507991206000704

CB Insights. "From virtual nurses to drug discovery: 106 artificial intelligence startups in healthcare." *CB Insights – Blog* (2017). Accessed 20 June 2017.

Cheng, Jie-Zhi, Dong Ni, Yi-Hong Chou, Jing Qin, Chui-Mei Tiu, Yeun-Chung Chang, Chiun-Sheng Huang, Dinggang Shen, and Chung-Ming Chen. "Computer-aided diagnosis with deep learning architecture: Applications to breast lesions in US images and pulmonary nodules in CT scans." *Scientific reports* 6 (2016): 244-254. doi: http://dx.doi.org/10.1038/srep24454

Collier, M., R. Fu, and L. Yin. "Artificial intelligence: Healthcare's new nervous system." *Accenture* (2017). Accessed 20 June 2017.

Delano-Wood, Lisa, Mark W. Bondi, Joshua Sacco, Norm Abeles, Amy J. Jak, David J. Libon, and Andrea Bozoki. "Heterogeneity in mild cognitive impairment: Differences in neuropsychological profile and associated white matter lesion pathology." *Journal of the International Neuropsychological Society* 15.6 (2009): 906-914. doi: http://dx.doi.org/10.1017/S1355617709990257

Drachman, David A. "The amyloid hypothesis, time to move on: Amyloid is the downstream result, not cause, of Alzheimer's disease." *Alzheimer's & Dementia* 10.3 (2013): 372-380. doi: http://dx.doi.org/10.1016/j.jalz.2013.11.003

Edmonds, Emily C., L. Delano-Wood, A. J. Jak, Douglas Galasko, D. P. Salmon, and M. W. Bondi. "'Missed' MCI: False-negative errors based on conventional diagnostic criteria." *Journal of Alzheimer's Disease* 52 (2016): 685-691. doi: http://dx.doi.org/10.3233/JAD-150986

Edmonds, Emily C., Lisa Delano-Wood, Lindsay R. Clark, Amy J. Jak, Daniel A. Nation, Carrie R. McDonald, David J. Libon, Rhoda Au, Douglas Galasko, David P. Salmon, and Mark W. Bondi. "Susceptibility of the conventional criteria for mild cognitive impairment to false-positive diagnostic errors." *Alzheimer's & Dementia* 11.4 (2015): 415-424. doi: http://dx.doi.org/10.1016/j.jalz.2014.03.005

Erfani, S.M., S. Rajasegarar, S. Karunasekera, and C. Leckie. "High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning." *Pattern Recognition* 58 (2016): 121.

Ertosun, M. G., and D. L. Rubin. 2015. "Automated grading of gliomas using deep learning in digital pathology images: A modular approach with ensemble of convolutional neural networks." *AMIA Annual Symposium Proceedings* (2015): 1899-908.

Esteva, Andre, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, and Sebastian Thrun. "Dermatologist-level classification of skin cancer with deep neural networks." *Nature* 542.7639 (2017): 115-118. doi: http://dx.doi.org/10.1038/nature21056

Finkel, Deborah, Marie Ernsth-Bravell, and Nancy L. Pedersen. "Sex differences in genetic and environmental influences on longitudinal change in functional ability in late adulthood." *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 70.5 (2015): 709-717. doi: http://dx.doi.org/10.1093/geronb/gbt134

Hardy, John, and Bart De Strooper. "Alzheimer's disease: where next for anti-amyloid therapies?" *Brain* 140.4 (2017): 853-855. doi: http://dx.doi.org/10.1093/brain/awx059

Klekociuk, S. Z., N. L. Saunders, and M. J. Summers. "Diagnosing mild cognitive impairment as a precursor to dementia: Fact or fallacy?" *Australian Psychologist* 51.5 (2016): 366-373. doi: http://dx.doi.org/10.1111/ap.12178

Klekociuk, Shannon Zofia, Jeffery J. Summers, James C. Vickers, and Mathew J. Summers. "Reducing false positive diagnosis in MCI: The importance of comprehensive neuropsychological assessment." *European Journal of Neurology* 21 (2014): 1330-1336. doi: http://dx.doi.org/10.1111/ene.12488

Korst, L.M., A.C. Eusebio-Angeja, T. Chamorro, C.E. Aydin, and K.D. Gregory. "Nursing documentation time during implementation of an electronic medical record." In *Evaluating the Organizational Impact of Healthcare Information Systems*, edited by J.G. Anderson and C.E. Aydin, 304-314. New York: Springer, 2005.

Littlefield, Christine H, John L Craven, Gary M Rodin, Denis Daneman, Michael A Murray, and Anne C Rydall. "Relationship of self-efficacy and binging to adherence to diabetes regimen among adolescents." *Diabetes Care* 15.1 (1992): 90-94. doi: http://dx.doi.org/10.2337/diacare.15.1.90

Madl, T. "Network analysis of heart beat intervals using horizontal visibility graphs." In *2016 Computing in Cardiology Conference (CinC)*, 11-14 Sept. 2016.

Madl, T. "Deep neural heart rate variability analysis." In *Neural Information Processing Systems – Machine Learning for Healthcare Barcelona, Spain* (2017).

Madl, Tamas, Stan Franklin, Ke Chen, Daniela Montaldi, and Robert Trappl. "Towards real-world capable spatial memory in the LIDA cognitive architecture." *Biologically Inspired Cognitive Architectures* 16 (2016): 87-104. doi: http://dx.doi.org/10.1016/j.bica.2016.02.001

Madl, Tamas, Stan Franklin, Ke Chen, Robert Trappl, and Daniela Montaldi. "Exploring the structure of spatial representations." *PLOS ONE* 11.6 (2016): e0157343. doi: http://dx.doi.org/10.1371/journal.pone.0157343

MarketsandMarkets Analysis. "Artificial Intelligence in Healthcare Market by Offering (Hardware, Software and Services), Technology (Deep Learning, Querying Method, NLP, and Context Aware Processing), Application, End-User Industry, and Geography – Global Forecast to 2022." 2017. Available at http://marketsandmarkets.com.

McKhann, Guy M., David S. Knopman, Howard Chertkow, Bradley T. Hyman, Clifford R. Jack, Claudia H. Kawas, William E. Klunk, Walter J. Koroshetz, Jennifer J. Manly, Richard Mayeux, Richard C. Mohs, John C. Morris, Martin N. Rossor, Philip Scheltens, Maria C. Carrillo, Bill Thies, Sandra Weintraub, and Creighton H. Phelps. "The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease." *Alzheimer's & Dementia* 7.3 (2011): 263-269.

Mufson, Elliott J., Michael Malek-Ahmadi, Sylvia E. Perez, and Kewei Chen. "Braak staging, plaque pathology, and APOE status in elderly persons without cognitive impairment." *Neurobiology of Aging* 37 (2016): 147-153. doi: https://doi.org/10.1016/j.neurobiolaging.2015.10.012

Neill, Denise. "Nursing workload and changing the health care environment." *Adminstrative Issues Journal: Education, Practice and Research* 1.2 (2011): 132-143.

Petersen, R.C., Glenn E. Smith, Stephen C. Waring, Robert J. Ivnik, Eric G. Tangalos, and Emre Kokmen. "Mild cognitive impairment: Clinical characterization and outcome." *Archives of Neurology* 56.3 (1999): 303-308.

Poissant, L., J. Pereira, R. Tamblyn, and Y. Kawasumi. "The impact of electronic health records on time efficiency of physicians and nurses: A systematic review." *Journal of the American Medical Informatics Association* 12.5 (2005): 505-516. doi: http://dx.doi.org/10.1197/jamia.M1700

Sarraf, S., G. Tofighi, and Alzheimer Disease Neuroimaging Initiative. "Deep learning-based pipeline to recognize Alzheimer's disease using fMRI data." In *2016 Future Technologies Conference (FTC)*, 6-7 Dec. 2016.

Suk, Heung-Il, Seong-Whan Lee, and Dinggang Shen. "Deep ensemble learning of sparse regression models for brain disease diagnosis." *Medical Image Analysis* 37 (2017): 101-113. doi: https://doi.org/10.1016/j.media.2017.01.008

Summers, Mathew J., and Nichole L. J. Saunders. "Neuropsychological measures predict decline to Alzheimer's dementia from mild cognitive impairment." *Neuropsychology* 26.4 (2012):498-508. doi: http://dx.doi.org/10.1037/a0028576

Varadhan, R., P. H. Chaves, L. A. Lipsitz, P. K. Stein, J. Tian, B. G. Windham, R. D. Berger, and L. P. Fried. "Frailty and impaired cardiac autonomic control: new insights from principal components aggregation of traditional heart rate variability indices." *The Journals of Gerontology: Series A* 64A.6 (June 2009): 682-687. doi: http://dx.doi.org/10.1093/gerona/glp013

Woo, Choong-Wan, Luke J. Chang, Martin A. Lindquist, and Tor D. Wager. "Building better biomarkers: brain models in translational neuroimaging." *Nat Neurosci* 20.3 (2017): 365-377. doi: http://dx.doi.org/10.1038/nn.4478

Ylikoski, Raija, Ari Ylikoski, Pertti Keskivaara, Reijo Tilvis, Raimo Sulkava, and Timo Erkinjuntti. "Heterogeneity of cognitive profiles in aging: Successful aging, normal aging, and individuals at risks for cognitive decline." *European Journal of Neurology* 6.6 (1999): 645-652.

Yu, Kun-Hsing, Ce Zhang, Gerald J. Berry, Russ B. Altman, Christopher Ré, Daniel L. Rubin, and Michael Snyder. "Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features." *Nature Communications* 7 (2016). doi: http://dx.doi.org/10.1038/ncomms12474