BJPsych | The British Journal of Psychiatry (2022)
221, 435–436. doi: 10.1192/bjp.2022.75

Kaleidoscope

Derek K. Tracy, Dan W. Joyce,
Dawn N. Albertson, Sukhwinder S. Shergill

History teaches us caution about 'dementia breakthroughs', but a fascinating paper in *Molecular Psychiatry* offers a development towards an eventual vaccine against Alzheimer's disease. Alzheimer's disease is characterised by deposits of the protein amyloid-beta (Aβ) that creates extracellular amyloid plaques in the brain. The influential cascade hypothesis suggests that this in turn leads to tau phosphorylation, tangle formation and ultimately cell death. There are various forms of the Aβ peptide, with much research exploring how these might differentially affect neurodegeneration. In particular, there is increasing evidence that soluble, non-plaque types of Aβ may offer better therapeutic targets than plaques, given their roles in synaptic plasticity and in inhibiting long-term potentiation. Bakrania et al[1] used TAP01, an antibody known to bind to non-plaque-forming Aβ, to identify a novel and unique conformational change – a pseudo β-hairpin structure in the N-terminal region of Aβ – that appears to be responsible for its binding properties.

From this, they engineered a form of Aβ1-14, 'TAPAS', with which they actively immunised two mouse models of Alzheimer's disease. This targeted the early toxic species of Aβ ubiquitous in Alzheimer's disease patients and led to a marked reduction in plaque formation (rather than reacting with the plaques themselves), rescue of brain glucose metabolism, stabilisation of neuron loss and recovery of some memory loss. A subsequent test with a humanised version of the TAP01 antibody produced similar effects. We know what you're thinking – #JustSaysInMice, which takes us back to the opening sentence on caution. Previous research on another putative vaccine had to be halted as it induced meningoencephalitis in a minority; further, although post-mortem examination showed greater plaque clearance, this did not prevent progressive neurodegeneration. Nevertheless, another piece of the pathogenesis of Alzheimer's disease appears to have become a little clearer, and the science has advanced a little further.

Frontotemporal dementias (FTDs) are devastating conditions, and one area of research is detecting early or prodromal states. FTDs are the fourth most common type of dementia and are notable for much earlier onset and the relatively strong autosomal dominant inheritance that occurs in up to a third of sufferers. There are three main subtypes: the behavioural variant (bvFTD), non-fluent or primary progressive aphasia and semantic variant primary progressive aphasia. Baker et al[2] explored bvFTD with 72 participants deemed to have a prodromal state based on clinician assessment and either being a carrier of a mutation known to cause FTD ($N = 55$) or autopsy confirmation ($N = 17$). A subgroup considered to have the strongest diagnostic evidence was evaluated in more detail to develop criteria for 'mild behavioural and/or cognitive impairment in bvFTD' (MBCI-FTD), which was then tested on the remaining validation group and familial non-carriers who acted as healthy controls.

Seven core features were identified: apathy without dysphoria, behavioural disinhibition, irritability/agitation, reduced empathy/sympathy, simple and/or complex repetitive behaviours, joviality/gregariousness, and appetite changes/hyperorality. The authors emphasised the presence of apathy *without* low mood – as depression could otherwise be a differential diagnosis or confounder – and the particular frequency of rapid agitation and anger, with an absence of insight. Although hallucinations and delusions did not appear to be prodromal features, the prodrome was associated with what the authors labelled 'supportive' features of impaired executive functioning with intact orientation and visuospatial skills, reduced insight into cognitive and behavioural changes, and poor social cognition. The authors propose that three core features, or two core and one supportive, are required for a diagnosis of *possible* MBCI-FTD, with *probable* needing additional imaging or biomarker evidence or a known pathogenic genetic mutation. These new criteria, the first for prodromal bvFTD, had a 95% diagnostic accuracy in the development group, with false positive rates of under 10% in healthy controls and 11–16% in those with prodromal Alzheimer's disease. Unfortunately we do not currently have good treatments for FTDs, but as these emerge, early identification will be helpful.

Young brains and the impact of environmental toxins on development. With increasing ageing of infrastructure, disadvantaged communities around the world are at increased risk of lead exposure from paint and environmental contamination. Lead exposure has a well-established relationship with poorer neurodevelopmental trajectories that continue into adulthood, with spiralling consequences across mental health and behaviour. Even following low levels of blood accumulation, adults with lead exposure as children have increased rates of criminality and lower socioeconomic mobility. Although the correlations among lead, adversity, and behaviour are strong, there is limited research investigating the interplay of all three factors. The Healthy Brains and Behavior Study[3] gathered adolescent children aged 11–12 years in the Philadelphia area of the USA. In addition to blood samples, children and their carers answered questions on a series of measures examining externalising and internalising behaviours, as well as a composite measure of social adversity, which included neighbourhood disadvantage and indicators around family structure, education and mental health. Higher levels of social adversity were associated with higher blood lead levels as well as externalising, but not internalising, factors. Ordinary least squares regression modelling revealed lead levels to be a mediating factor between social adversity and externalising behaviour. Importantly, the effect was seen at blood levels around half of those usually used in this type of research.

Measuring social adversity with multiple indicators beyond socioeconomic status helps to more fully capture the cumulative stressors with negative impact on children's development. Although further work will be necessary to articulate the specific influences of lead exposure and its accumulation across developmental timelines, the message is clear. The impact of even the trace amounts of lead more commonly seen in the global north necessitate a more aggressive approach to elimination as a public health priority worldwide. In the meantime, we must reckon with the impact that environmental factors like this can have on child and adolescent mental health and how we may take this into consideration when dealing with the behavioural outcomes in schools and juvenile reform systems.

The *BJPsych* recently published[4] its (disappointing) author gender data; has there been improvement across other high-impact journals? Krstacic et al[5] evaluated the heavy hitters of the *New England Journal of Medicine* (*NEJM*), the *Journal of the American Medical Association* (*JAMA*), and the *Lancet*. They looked at the first, second and last authors, extracting gender based on internet biographies, pronouns, names and photographs. Considering almost two decades, from 2002 to 2019, the authors also explored any trends across time. Overall, women accounted for approximately 16%, 35% and 29% of first authors in *NEJM*, *JAMA* and the *Lancet*, contrasted against 37% of US medical school full-time faculty during the same period. The data were no

better for second or last author, and, really disappointingly, there has been no improvement over time. The authors note how women authors also had lower citation counts, fewer multiple publications in these journals and fewer multiple doctoral-level degrees. None of this is explicitly blaming higher-impact journals; it's just symptomatic of what is happening everywhere and, certainly, life is no better at the academic peak. Something has to change, and just measuring data does not seem to be moving the needle (although the *BJPsych* has committed to continue to publish its own figures on gender). A broader system-wide approach is needed: research funders, universities and scientific journals need to have some hard targets set to turn this dial.

Finally, the management guru Peter Drucker said 'Trying to predict the future is like trying to drive down a country road at night with no lights while looking out the back window'. In many ways, psychiatry's fascination with predictive analytics reflects this. Whether we use machine learning, techniques from artificial intelligence (more generally) or our more trusted statistical explanatory/inferential models, our country road is the complex nosology and aetiology of mental illness, the lack of lighting reflects our use of opaque black-box software toolkits, and looking out the back window equates to not paying enough attention to what we (and these complex tools) are doing. To help us navigate better, Meehan et al[6] systematically reviewed around 20 years of clinical prediction models – from a total of at 308 studies – evaluating each against benchmark audit tools for quality and readiness for clinical use. Specifically, they examined each study for: minimisation of bias or systematic error and overfitting; sample size and sufficient data for the outcome being predicted or discriminated; generalisability (i.e. the performance of the model in 'unseen' data or individuals) and clinical utility (i.e. when compared with existing prediction rules, did the model add value above current practice or models). Overall, 215 (just under 70%) were prognostic models (e.g. onset of illness or some future outcome of interest), 66 (21%) were diagnostic (e.g. identifying cases) and 27 were predictive. In terms of disease areas, outcomes in depression was the most studied, followed by psychosis and then post-traumatic stress disorder. Most of the literature (291 of 308 papers) reported model development – that is, testing the ability to build a multivariable model for some disease areas and outcome – but only 70 papers included or explicitly examined external validation (the most valuable test of generalisability).

In summary, 283 of 308 papers were at high risk of bias because there were insufficient outcome events in the data relative to the total number of included predictors or model parameters. That is, the event of interest was too infrequently present in the data to allow robust modelling. Another potential risk of bias (in model development, sometimes called 'model training') is the use of automated (data-driven) variable selection methods where one 'throws' all variables into an algorithm that shrinks the total number of predictor variables in the final model to some subset deemed 'relevant'. Some of these methods are known to lead to biased model performance because of how they define a 'relevant' predictor. In particular, forward/backward and stepwise selection methods were used in 59 of 220 model development papers (26%), with bivariate correlations between predictors and outcome variables representing 54/220 (24%) – methods that are generally not recommended. In 64 (of all 308) studies, models were selected by comparing many statistical or machine learning methods and then choosing the 'best' model based on its performance on the data-set. Ultimately, we need to know how well a model performs to assess its utility. For predictive models which give us a 'yes/no' answer (i.e. the patient has depression or not), discrimination performance is given by measures such as the area under the receiver operating curve; or, for summary point-estimates such as sensitivity, specificity and accuracy, one needs to know the threshold at which a positive or negative decision was declared by the model. For models making continuous-valued, forecast-like predictions (i.e. the probability of, time to, or risk of an event), calibration is the better measure – the minimum requirement being a visualisation showing the continuous-valued model output (on the horizontal axis) against the event of interest (on the vertical axis), where a 'good' calibration would be indicated by an approximately straight diagonal line. Importantly, one should report the discrimination and/or calibration as required by the type of model being developed. Meehan et al found that discrimination metrics were reported in 271/308 papers and calibration in 68/308. Of those reporting point-estimates of discrimination (38 of 308 papers), only nine reported the threshold used for declaring positive/negative outputs. Finally, to clinical utility, which evaluates whether the predictive model 'adds' value compared with usual practice. Decision curve analysis demonstrates the 'net benefit' of a model and is the most commonly used method – but only two papers (of 308) attempted this. Often, a nascent field begins with well-intentioned excitement, resulting in a flurry of papers offering new understanding; a mature field is one that reflects and adjusts its path in the light of 'debugging' its literature to ensure it develops responsibly and usefully.

References

1 Bakrania P, Hall G, Bouter Y, Beindorff N, Cowan R, Davies S, et al. Discovery of a novel pseudo β-hairpin structure of N-truncated amyloid-β for use as a vaccine against Alzheimer's disease. *Mol Psychiatry* 2022; **27**: 840–8.

2 Barker MS, Gottesman RT, Manoochehri M, Chapman S, Appleby BS, Brushaber D, et al. Proposed research criteria for prodromal behavioural variant frontotemporal dementia. *Brain* 2022; **145**(3): 1079–97.

3 Liu J, Portnoy J, Raine A, Gladieux M, McGarry P, Chen A. Blood lead levels mediate the relationship between social adversity and child externalizing behavior. *Environ Res* 2022; **204**(Pt D): 112396.

4 Upthegrove R, de Cates A, Shuttleworth A, Tracy DK, Broome MR, Lingford-Hughes A. Gender equality in academic publishing: action from the *BJPsych*. *Br J Psychiatry* 2021; **218**(3): 128–30.

5 Krstacic JE, Carr BM, Yaligar AR, Kuruvilla AS, Helali JS, Saragossi J, et al. Academic medicine's glass ceiling: author's gender in top three medical research journals impacts probability of future publication success. *PLoS One* 2022; **17**(4): e0261209.

6 Meehan AJ, Lewis SJ, Fazel S, Fusar-Poli P, Steyerberg EW, Stahl D, et al. Clinical prediction models in psychiatry: a systematic review of two decades of progress and challenges. *Mol Psychiatry* [Epub ahead of print] 1 Apr 2022. Available from: https://doi.org/10.1038/s41380-022-01528-4.

https://doi.org/10.1192/bjp.2022.75 Published online by Cambridge University Press