



## LJMU Research Online

**Bruno, D and Jauregi Zinkunegi, A**

**The recency ratio assessed by story recall is associated with cerebrospinal fluid levels of neurodegeneration biomarkers**

<http://researchonline.ljmu.ac.uk/id/eprint/18362/>

### Article

**Citation** (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

**Bruno, D and Jauregi Zinkunegi, A (2023) The recency ratio assessed by story recall is associated with cerebrospinal fluid levels of neurodegeneration biomarkers. Cortex, 159. pp. 167-174. ISSN 0010-9452**

LJMU has developed **LJMU Research Online** for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact [researchonline@ljmu.ac.uk](mailto:researchonline@ljmu.ac.uk)

<http://researchonline.ljmu.ac.uk/>



## Behavioural Neurology

# The recency ratio assessed by story recall is associated with cerebrospinal fluid levels of neurodegeneration biomarkers

Davide Bruno <sup>a,\*</sup>, Ainara Jauregi Zinkunegi <sup>a</sup>, Gwendlyn Kollmorgen <sup>b</sup>,  
 Ivonne Suridjan <sup>c</sup>, Norbert Wild <sup>b</sup>, Cynthia Carlsson <sup>d,e,f,g</sup>,  
 Barbara Bendlin <sup>e,f</sup>, Ozioma Okonkwo <sup>e,f</sup>, Nathaniel Chin <sup>d,e</sup>,  
 Bruce P. Hermann <sup>d,h</sup>, Sanjay Asthana <sup>d,e</sup>, Henrik Zetterberg <sup>i,j,k,l,m</sup>,  
 Kaj Blennow <sup>i,j</sup>, Rebecca Langhough <sup>d,e,f</sup>, Sterling C. Johnson <sup>d,e,f,g</sup> and  
 Kimberly D. Mueller <sup>d,e,n</sup>

<sup>a</sup> School of Psychology, Liverpool John Moores University, UK

<sup>b</sup> Roche Diagnostics GmbH, Penzberg, Germany

<sup>c</sup> Roche Diagnostics International Ltd, Rotkreuz, Switzerland

<sup>d</sup> Wisconsin Alzheimer's Institute, School of Medicine and Public Health, University of Wisconsin – Madison, Madison, WI, USA

<sup>e</sup> Wisconsin Alzheimer's Disease Research Center, School of Medicine and Public Health, University of Wisconsin – Madison, Madison, WI, USA

<sup>f</sup> Department of Medicine, University of Wisconsin-Madison, Madison, WI, USA

<sup>g</sup> Geriatric Research Education and Clinical Center, William S. Middleton Veterans Hospital, Madison, WI, USA

<sup>h</sup> Department of Neurology, University of Wisconsin – Madison, Madison, WI, USA

<sup>i</sup> Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, The Sahlgrenska Academy at the University of Gothenburg, Mölndal, Sweden

<sup>j</sup> Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden

<sup>k</sup> Department of Neurodegenerative Disease, UCL Institute of Neurology, Queen Square, London, UK

<sup>l</sup> UK Dementia Research Institute at UCL, London, UK

<sup>m</sup> Hong Kong Center for Neurodegenerative Diseases, Clear Water Bay, Hong Kong, China

<sup>n</sup> Department of Communication Sciences and Disorders, University of Wisconsin – Madison, Madison, WI, USA

## ARTICLE INFO

## Article history:

Received 19 July 2022

Reviewed 3 November 2022

Revised 11 November 2022

Accepted 8 December 2022

Action editor Stefano Cappa

Published online 23 December 2022

## ABSTRACT

Recency refers to the information learned at the end of a study list or task. Recency forgetting, as tracked by the ratio between recency recall in immediate and delayed conditions, i.e., the recency ratio (Rr), has been applied to list-learning tasks, demonstrating its efficacy in predicting cognitive decline, conversion to mild cognitive impairment (MCI), and cerebrospinal fluid (CSF) biomarkers of neurodegeneration. However, little is known as to whether Rr can be effectively applied to story recall tasks. To address this question, data were extracted from the database of the Alzheimer's Disease Research Center at the University of Wisconsin – Madison. A total of 212 participants were included in the study. CSF

\* Corresponding author. Tom Reilly Building, Byrom Street, Liverpool, L3 3AF, UK.

E-mail address: [d.bruno@ljamu.ac.uk](mailto:d.bruno@ljamu.ac.uk) (D. Bruno).

<https://doi.org/10.1016/j.cortex.2022.12.004>

0010-9452/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

**Keywords:**

Memory assessment  
 Story recall  
 Dementia  
 Serial position effects  
 Cerebrospinal fluid biomarkers

biomarkers were amyloid-beta ( $A\beta$ ) 40 and 42, phosphorylated (p) and total (t) tau, neurofilament light (NFL), neurogranin (Ng), and  $\alpha$ -synuclein (a-syn). Story Recall was measured with the Logical Memory Test (LMT). We carried out Bayesian regression analyses with Rr, and other LMT scores as predictors; and CSF biomarkers (including the  $A\beta$ 42/40 and p-tau/ $A\beta$ 42 ratios) as outcomes. Results showed that models including Rr consistently provided best fits with the data, with few exceptions. These findings demonstrate the applicability of Rr to story recall and its sensitivity to CSF biomarkers of neurodegeneration, and encourage its inclusion when evaluating risk of neurodegeneration with story recall.

© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In memory research, serial position effects refer to better retrieval of the information learned at the beginning of a study list or task as compared to the information learned in the middle (primacy effect), and better retrieval of the information learned at the end of such study list or task as compared to the information learned in the middle (recency effect; [Murdock, 1962](#)). Interestingly, while individuals with Alzheimer's disease (AD) typically present with poor primacy effects, a recency effect is often still observed, particularly when testing occurs right after the learning phase ([Foldi et al., 2003](#)). However, recency performance then tends to deteriorate if a delay is placed between learning and test ([Carlesimo et al., 1995](#)). To leverage this pattern whereby high recency is observed at immediate recall whereas low recency is found at delayed recall in AD, the recency ratio (Rr) was proposed, which is operationally defined as the ratio between recency recall in immediate compared to delayed recall conditions ([Bruno et al., 2016, 2018](#)).

Using Rr in list-learning tasks (e.g., Rey's Auditory Verbal Learning Task), it has been shown that higher (i.e., worse) scores predict cognitive decline in asymptomatic individuals ([Bruno et al., 2016](#)), preclinical mild cognitive impairment (MCI) from a healthy baseline ([Bruno et al., 2016; Egeland, 2021](#)), and amyloid- $\beta$  pathology in individuals with MCI ([Bruno et al., 2019](#)). Rr from list-learning tasks has also been found to correlate with cerebrospinal fluid (CSF) levels of neurogranin (Ng), a post-synaptic protein reflecting synaptic dysfunction ([Bruno, Reichert Plaska, et al., 2021](#)), and to both phosphorylated (p-) and total (t-) tau levels ([Bruno et al., 2022](#)). Additionally, Rr scores have been found to discriminate between individuals diagnosed with AD versus other types of dementia ([Turchetta et al., 2018](#)), and to identify successfully individuals with MCI who are more likely to convert to AD ([Turchetta et al., 2020](#)). All in all, in list-learning tasks, Rr compares favourably to most conventional scores employed to estimate memory ability in older individuals ([Bock et al., 2021](#)).

However, little is known as to how successful Rr may be in story recall tasks. Story recall tasks, unlike list-learning tasks, present participants with a coherent story to learn, and then typically ask them to recall it right after presentation and then again after a delay. Owing to their semantic structure

(compared to a semantically unrelated list of words), story recall tasks are typically thought to be less sensitive to serial position effects, but this is not the case. [Hall and Bornstein \(1991\)](#), for example, showed serial position effects in individuals with closed head injuries and controls, when using the Logical Memory test (LMT), a common story recall task ([Wechsler, 1987](#)). The LMT is also commonly used as a screening tool for dementia, as demonstrated by its use in the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the Australian Imaging, Biomarker & Lifestyle Flagship Study of Ageing (AIBL). Aptly then, [Bruno, Mueller, et al. \(2021\)](#) examined LMT scores in late middle-age and older individuals with MCI and people with no cognitive impairment, and observed clear serial position effects. Moreover, they also highlighted the role of primacy forgetting (billed there *primacy ratio*, henceforth Pr) in predicting amyloid burden.

The aims of the present study were to examine the degree to which process scores such as Rr and Pr may predict CSF biomarkers of brain amyloidosis, tau pathology and neurodegeneration in middle age and older individuals. Moreover, we also aimed to establish whether these process scores performed better than traditional scores of immediate and delayed recall performance in story recall. We achieved this by analysing data from the Wisconsin Alzheimer's Disease Research Center (WADRC), an ongoing longitudinal cohort study based at the University of Wisconsin – Madison, which included LMT data alongside CSF levels of biomarkers associated with AD and neurodegeneration: amyloid-beta ( $A\beta$ ) 40 and 42, p- and t-tau, neurofilament light (NFL), Ng, and  $\alpha$ -synuclein (a-syn).

## 2. Methods

We report how we determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study. The ethical regulations that govern the WADRC prevent unrestricted public archiving of anonymised study data. Data can be requested from the WADRC Executive Committee at: <https://www.adrc.wisc.edu/apply-resources>. Data will be released to internal and external investigators following confirmation of IRB approval together with an evaluation by the WADRC of scientific merit and resource availability.

## 2.1. Participants

Data were extracted from the WADRC database. After baseline, WADRC participants complete regular follow-up visits at 1- or 2-year intervals, including neuropsychological tests, clinical measurements (e.g., blood pressure, heart rate), and health history. Participants were selected on the basis of having completed at least two assessment visits: one for cognitive evaluation, including LMT, and at least one lumbar puncture visit for CSF extraction. Participants were classified as either cognitively unimpaired, with MCI due to presumed AD (MCI-AD), or with dementia due to presumed AD, via consensus conference diagnosis, determined by a team that included physicians, clinical neuropsychologists, and clinical nurse practitioners, based on core clinical criteria developed by the National Institute on Aging and the Alzheimer's Association (Albert et al., 2011; McKhann et al., 2011), and without regard to AD biomarker status. From the total pool of 828 participants, 212 participants fit the inclusion criteria. Of these, 156 were cognitively unimpaired, 26 had a diagnosis of MCI-AD, and 30 had a diagnosis of AD. Moreover, 16 (10%) cognitively unimpaired participants displayed biomarkers-determined AD based on a CSF p-tau/A $\beta$ 42 ratio cut-off of .038: a low ratio identifies individuals without a biomarker-based AD diagnosis, whereas a higher ratio identifies people with a biomarker-based AD diagnosis (Van Hulle et al., 2021). Similarly, 16 (62%) and 28 (93%) of MCI-AD and AD participants, respectively, displayed AD biomarkers. Only one participant did not identify as white Caucasian. Participants' cognitive data were taken from whichever visit was closest to the visit where the lumbar puncture was performed. All activities for this study were approved by the ethics committees of the authors' universities and competed in accordance with the Declaration of Helsinki. All participants provided informed consent prior to testing. No part of the study procedures or analyses were pre-registered prior to the research being conducted.

## 2.2. Cognitive assessment

The LMT was used to measure story recall performance. The LMT is a subtest of the Wechsler Memory Scale Revised (WMS-R; Wechsler, 1987), comprising two stories, A and B, with 25 items (“idea units”) each, and representing different semantic and lexical categories. Each story is read aloud to the participant and then the participant is asked to recall both stories immediately and again after a 25–30-min delay. Scoring procedures from the WMS-R manual were applied. Although the scoring criteria permits some alteration from the original item (e.g., “slid off the table” is allowed instead of “fell off the table”), certain items must be recalled verbatim, e.g., numerical expressions or proper names. In order to slim down the administered cognitive battery, story recall in the present study was measured only with story A of the LMT. However, to note, story A and story B have been found to be of comparable memorability in a separate study (Mueller et al., 2022). Immediate LMT and Delayed LMT recall scores were calculated from adding all the correctly recalled items in the immediate recall trial and delayed recall trial, respectively. Possible scores for Immediate and Delayed Recall trials range from 0 to

25 for each, where higher scores reflect more items being recalled. Finally, primacy and recency were defined as the first and final eight idea units of the story, respectively, while middle was defined as the middle nine units, following previous work (Bruno, Mueller, et al., 2021) – the choice of idea units per serial position is arbitrary. Immediate and delayed recency scores were calculated as the number of correctly recalled recency items in immediate and delayed recall trials, respectively. R<sub>r</sub> was obtained by dividing the recency scores in the immediate recall trial by the corresponding scores in the delayed recall trial. A correction also was applied ((immediate recency score + 1)/(delayed recency score + 1)) to avoid missing data due to zero scores (Bruno et al., 2018). P<sub>r</sub> was calculated following (Bruno, Mueller, et al., 2021) by dividing delayed primacy by immediate primacy, with no adjustments. While this is inconsistent with the way R<sub>r</sub> is computed, and should arguably be aligned, we opted here for maintaining the original formula. Finally, to provide a non-serial position based forgetting index that would account for memory loss, we also computed a ratio score with Immediate LMT and Delayed LMT ((Immediate LMT + 1)/(Delayed LMT + 1)), which we dubbed the total ratio (Tr).

## 2.3. Biomarker determination

All CSF samples were assayed at the Clinical Neurochemistry Laboratory, University of Gothenburg, under strict quality control procedures. All CSF markers were measured using the exploratory Roche NeuroToolKit assays, a panel of automated robust prototype immunoassays (Roche Diagnostics International Ltd, Rotkreuz, Switzerland), not currently approved for clinical use. Elecsys® A $\beta$ 42, A $\beta$ 40, p-tau (181P), and t-tau, were performed on a cobas e 601 analyzer; a-syn, NFL, and Ng were performed on a cobas e 411 analyzer, as previously described (Van Hulle et al., 2021).

## 2.4. Genotyping

DNA was extracted from whole blood. Samples were aliquoted on 96-well plates for determination of APOE genotypes. An APOE risk score was calculated based on the odds ratios of the e2/e3/e4 genotype, as previously reported (Darst et al., 2017).

## 2.5. Analysis plan

For each CSF outcome, we carried out Bayesian linear regressions with P<sub>r</sub>, R<sub>r</sub>, Immediate LMT, Delayed LMT, and Tr as predictors in the same analyses; age at the lumbar puncture, time elapsed between lumbar puncture and memory assessment, sex, years of education, APOE risk score, number of overall cognitive assessment visits (to account for practice effects), and whether they classified for biomarkers-determined AD were used as control variables. Control variables formed the null models. CSF biomarkers were used as outcomes, in separate analyses. Bayesian analyses allow for the estimation of model plausibility, which permits comparison of models with different combinations of predictors, and for the determination of effect sizes with credible intervals (e.g., Teipel et al., 2021). In [Supplementary Materials](#) we also report the outputs of Frequentist analyses. Note that diagnosis

was not included in the analyses to avoid circularity since LMT scores are evaluated in the consensus process. For all Bayesian analyses, the model prior was set to Uniform, where all models are *a-priori* equally likely, and the prior on parameters was set to the default Jeffreys-Zellner-Siow (JZS) prior probability distribution, which allows the Bayes factor to be the same regardless of unit of measurement. Credible intervals were set to 95%. To address potential issues with non-normally distributed residuals in the regressions, Markov chain-Monte Carlo (MCMC) sampling to each analysis was applied 1,000 times. The outcome variables were CSF levels of Aβ40, Aβ42, p-tau, t-tau, Ng, NFL, and a-syn. Additionally, we also examined models with Aβ42/Aβ40 and p-tau/Aβ42 ratios as outcomes, as these measures are commonly used as biomarkers of neurodegeneration (Campbell et al., 2021; Li et al., 2013). Control variables formed the null models in each analysis. Analyses were conducted using JASP (0.16.2; <https://jasp-stats.org/>).

### 3. Results

Table 1 reports means, standard deviations and range for the demographic variables, age differences, APOE risk score, Pr, Rr, Immediate LMT, Delayed LMT, and Tr recall scores by cognitive status closest to lumbar puncture. Rr scores ranged from .43 to 5 across participants. Note that while higher Pr scores are preferable, the higher the Rr score, the worse.

Fig. 1a and b report serial position performance by delay in controls and individuals with biomarkers-determined AD, respectively. The values for primacy, middle and recency are proportions out of eight, nine and eight, respectively, to allow for direct comparison across serial positions. The plot displays a slightly more pronounced curve for immediate recall than for delayed recall in controls, and a substantial drop in

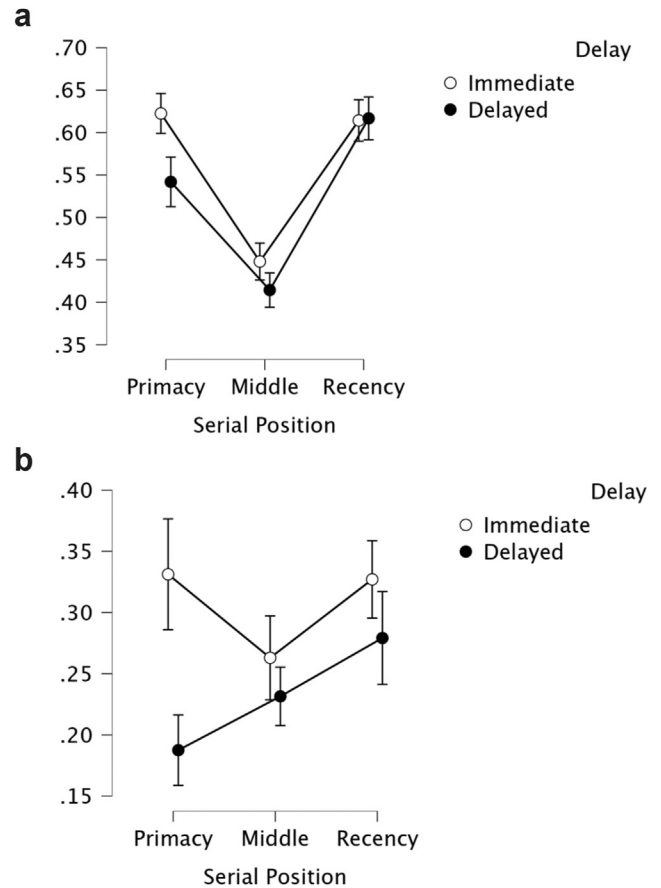


Fig. 1 – a. Serial position plot by delay in controls. Imm: immediate recall; Del: delayed recall. b. Serial position plot by delay in individuals with AD. Imm: immediate recall; Del: delayed recall.

Table 1 – Demographics, CSF measures and memory tests scores (mean and standard deviation) for the study participants. Elapsed time refers to time between cognitive testing and lumbar puncture, and it was calculated as an absolute value. Statistical tests were also conducted to check for differences across cognitively unimpaired, MCI-AD and AD: *p* values are reported. LP = lumbar puncture; CSF = cerebro-spinal fluid; Rr = recency ratio; Tr = total ratio; Pr = primacy ratio; LMT = logical memory test. \*Aβ40 N = 211, p-tau N = 212, a-syn N = 212.

Characteristic	Total (N = 212)	Cognitively unimpaired (N = 156)	MCI-AD (N = 26)	AD (N = 30)	<i>p</i> value
Sex (female)	129 (61%)	112 (72%)	6 (23%)	11 (37%)	<.001
Education (years)	16.1 (2.6)	16.2 (2.4)	16.5 (2.8)	14.8 (3.0)	.014
Age at LP (years)	65.1 (10.0)	62.1 (8.6)	74.6 (8.4)	72.7 (9.2)	<.001
Elapsed time (years)	.2 (.2)	.2 (.2)	.2 (.1)	.1 (.1)	.146
APOE risk score	1.4 (.9)	1.2 (.8)	1.6 (.8)	2.0 (1.1)	<.001
CSF Aβ42 (ng/L)	824.1 (390.8)	939.2 (376.7)	579.7 (214.6)	445.1 (199.3)	<.001
CSF Aβ40 (ng/L)*	14286.0 (4648.0)	14284.9 (4591.5)	14496.9 (4666.2)	14108.3 (5062.2)	.953
CSF Aβ42/Aβ40* (ng/L)	.1 (.02)	.06 (.0)	.04 (.0)	.03 (.0)	<.001
CSF P-tau (ng/L)	21.3 (12.8)	17.0 (6.9)	29.0 (18.2)	36.7 (15.3)	<.001
CSF P-tau/Aβ42 (ng/L)	.04 (.04)	.02 (.0)	.06 (.0)	.09 (.0)	<.001
CSF T-tau (ng/L)	232.8 (116.9)	195.1 (76.4)	301.2 (155.5)	366.8 (128.9)	<.001
CSF α-synuclein* (ng/L)	172.7 (84.8)	155.4 (70.8)	206.4 (107.8)	232.2 (95.4)	<.001
CSF NFL (ng/L)	120.1 (110.6)	91.5 (79.5)	179.6 (153.4)	215.5 (133.0)	<.001
CSF Ng (ng/L)	818.9 (363.0)	753.6 (311.9)	917.8 (475.5)	1068.1 (382.1)	<.001
Rr	1.1 (.5)	1.0 (.2)	1.3 (.8)	1.6 (.9)	<.001
Immediate LMT	12.1 (5.2)	14.5 (3.2)	7.9 (4.9)	3.9 (2.8)	<.001
Delayed LMT	11.0 (5.6)	13.6 (3.3)	5.4 (4.7)	2.0 (2.1)	<.001
Tr	1.3 (.9)	1.1 (.1)	2.0 (1.7)	2.0 (1.4)	<.001
Pr	.7 (.4)	.9 (.2)	.3 (.3)	.1 (.2)	<.001

delayed primacy in biomarkers-determined AD. This pattern is analogous to that reported already by Bruno, Mueller, et al. (2021). It may be noted also that delayed recency is better in our data with LMT than what traditionally expected with word-list tests. While it is beyond the scope of this paper to address theories of serial position, these findings do argue against the recency boost being solely a consequence of short-term memory processing in story recall.

**CSF A $\beta$ 42.** The best fitting model was the null model. The second best model had Immediate LMT alone (see Supplementary information for full model comparisons and posterior summaries). The Bayes Factor ( $BF_{10}$ ) that gives us the relative predictive adequacy of this model compared to the null model was .649, meaning that the observed data are .649 times more likely under this model than under the null model (which includes all the covariates).  $BF_{10}$  scores below 1, as in this case, indicate that the null model is a better fit for the data than the alternative models. Conventionally, also,  $BF_{10}$  scores below 3 are considered to provide only anecdotal evidence over the null model, and are therefore not sufficiently strong to draw firm conclusions.

**CSF A $\beta$ 40.** In contrast, the best fitting model for CSF A $\beta$ 40 combined Delayed LMT performance with Rr ( $BF_{10} = 5.350$ ; moderate evidence). Both Delayed LMT and Rr were positively associated with CSF A $\beta$ 40 levels: Delayed LMT had a posterior mean of 64.438 (SD = 112.988), and 95% Credible Intervals (CIs) ranged from -84.286 to 327.865; Rr had a posterior mean of 1067.043 (SD = 844.836), and CIs ranged from -47.522 to 2539.844. The inclusion probability was much higher for Rr, .762, than for Delayed LMT, .493, suggesting that Rr is the better predictor of CSF A $\beta$ 40 in these data.

**CSF t-tau.** The best fitting model with CSF t-tau included only Rr ( $BF_{10} = 3754.173$ ; extreme evidence). Rr had a posterior mean of 56.214 (SD = 13.652; CIs from 31.435 to 83.779). The inclusion probability for Rr was >.999, and adding Rr to the model improved it by over 650 times (i.e.  $BF_{inclusion} = 658.873$ ).

**CSF p-tau.** The best fitting model with CSF p-tau also included only Rr ( $BF_{10} = 1552.309$ ; extreme evidence). Rr had a posterior mean of 5.490 (SD = 1.428; CIs from 2.719 to 8.241). The inclusion probability for Rr was = .998 (i.e.  $BF_{inclusion} = 231.593$ ).

**CSF Ng.** The best fitting model with CSF Ng combined Delayed LMT and Rr ( $BF_{10} = 392.515$ ; extreme evidence). Again, the  $BF_{inclusion}$  for Rr (62.712) trumped that for delayed LMT (6.596). Higher Delayed LMT performance was, unexpectedly, associated with higher levels of CSF Ng (posterior mean = 21.180, SD = 17.387, CIs: 0 to 55.836), whereas as predicted Rr was positively correlated with Ng: posterior mean = 184.119, SD = 55.681, CIs: 84.228 to 299.589.

**CSF NFL.** With CSF NFL levels, the best fitting model included Immediate LMT, Tr and Rr ( $BF_{10} = 105.934$ ; extreme evidence). Immediate LMT was negatively associated with NFL (posterior mean = -4.492, SD = 3.327, CIs: -11.267 to 0), Tr also, and against expectations, was negatively correlated with NFL levels (posterior mean = -9.042, SD = 11.231, CIs: -34.594 to 0). In contrast, Rr was positively correlated with NFL (posterior mean = 29.297, SD = 21.281, CIs: 0 to 66.394), as expected. The inclusion probability was a little higher for Immediate LMT, .817,  $BF_{inclusion} = 4.192$ , compared to Rr, .790,  $BF_{inclusion} = 3.530$ , and Tr, .535,  $BF_{inclusion} = 1.225$ .

**CSF a-syn.** The best fitting model with CSF a-syn was the model with Rr alone ( $BF_{10} = 50.318$ ; very strong evidence). Rr was positively correlated with a-syn (posterior mean = 33.501, SD = 12.311, CIs: 12.009 to 59.343). The  $BF_{inclusion}$  was 28.196.

**CSF A $\beta$ 42/A $\beta$ 40.** The best fitting model was the null model, followed by a model with Rr alone ( $BF_{10} = .250$ ).

**CSF p-tau/A $\beta$ 42.** The best fitting model had Immediate LMT and Rr together ( $BF_{10} = 61,549.487$ ; extreme evidence). Immediate LMT had a posterior mean of -.001 (SD < .001; CIs -.003 to 0); and Rr had a posterior mean of .012 (SD = .003; CIs from .006 to .018). The best inclusion probability was for Rr, >.999 ( $BF_{inclusion} = 580.040$ ), whereas Immediate LMT reached .745 ( $BF_{inclusion} = 2.636$ ).

## 4. Discussion

The goal of this study was to establish whether process scores from story recall, such as Rr, were as sensitive to amyloid and tau proteinopathy, alongside other biomarkers of neurodegeneration, as it has been shown to be previously in list-learning tasks. To test this claim, we analysed data from the WADRC, comprising 212 participants who were either cognitively unimpaired, with presumed MCI-AD, or with presumed AD, and we correlated performance in the LMT, a popular story recall test, with CSF levels of biomarkers associated with AD, including measures of amyloid, tau, and neurodegeneration. Our Bayesian analyses clearly indicate that cross-sectional Rr levels are associated with several CSF biomarkers of neurodegeneration and AD, when controlling for age, level of education and APOE risk. Rr also specifically outperformed a ratio score introduced to measure total memory loss from immediate to delayed story recall, which we termed here total ratio. By and large, these results were mirrored by frequentist analyses (see [Supplementary information, S1](#)): Rr was the best predictor of t- and p-tau, Ng, a-syn and the p-tau/A $\beta$ 42 ratio, consistent with the Bayesian results. Also consistent with the Bayesian results, Rr was not correlated with A $\beta$ 42, NFL and the A $\beta$ 42/A $\beta$ 40 ratio. The only difference was with A $\beta$ 40, where the Bayesian analysis, but not the Frequentist analysis, found an association with Rr.

A lack of association between Rr and CSF A $\beta$ 42 is overall consistent with a recent report (within an overlapping cohort) using a list-learning task, where Rr was not found to predict CSF A $\beta$ 42 levels (Bruno et al., 2022). As CSF A $\beta$ 42 levels are thought to reflect closely brain amyloid deposition, the pre-eminent pathological hallmark of AD, these findings may suggest that Rr is not a specific cognitive marker of AD. However, this notion is not consistent with the following observations: Rr is sensitive to the CSF levels of Ng (also reported in Bruno, Reichert Plaska, et al., 2021), a neuron-specific postsynaptic protein that has been linked specifically to AD neurodegeneration (Wellington et al., 2016; Zetterberg & Bendlin, 2021); and Rr was also sensitive to the levels of the p-tau/A $\beta$ 42 ratio, which has been shown to be as predictive of brain amyloid pathology (Campbell et al., 2021). A final point to consider is that Rr was also found to correlate with CSF a-syn levels, partially consistent with the results of Bruno, Reichert Plaska, et al. (2021). While a-syn, a pre-synaptic protein that

can be found in cortical and sub-cortical areas, is typically linked to Parkinson's disease and dementia with Lewy bodies (Selnes et al., 2017), elevated CSF a-syn levels have also been found in individuals on a trajectory to AD (Shim et al., 2020).

Both p- and t-tau were found to be associated with Rr levels, suggesting that it is sensitive to neurofibrillary tangle (tau) pathology and, in turn, neurodegeneration in the medial-temporal lobe (Maass et al., 2019; Tennant et al., 2021). These findings are consistent with a recent report using list-learning, where Rr was also found to correlate positively with both p- and t-tau in individuals with MCI and unimpaired cognition (Bruno et al., 2022). A link between higher Rr scores and lower volume of the hippocampus was also recently observed (Jauregi et al., 2022) in overlapping participants, giving credence to the suggestion that higher Rr scores may be a consequence of combined loss of consolidation ability, which would follow atrophy of the medial-temporal lobe (Wixted, 2004; Wixted & Cai, 2013), while reliance on phonological/echoic short-term memory remains relatively intact (Bruno et al., 2018; Turchetta et al., 2018). These observation may also help explaining further the lack of association between Rr and CSF A $\beta$ 42 levels, as amyloid pathology does not specifically target regions in the medial temporal lobe, unlike neurofibrillary tangle pathology.

The different sex distribution across consensus diagnoses should be noted. As per Table 1, 72% of unimpaired individuals were female, while that percentage dropped drastically in people with MCI (23%) or probable AD (37%). This finding is at odds with the common observation that the majority of AD cases tend to be women (Alzheimer's Association, 2017). While a thorough examination of this issue is beyond the scope of the present manuscript, we looked at story recall outputs across sexes to see how they may vary. Interestingly, while immediate and delayed LMT scores are significantly higher for unimpaired females than males, with both parametric and non-parametric tests, Rr tends not to vary in relation to sex in this group. Finally, none of the memory scores differed across sexes for people with MCI or AD.

Despite co-varying in the analysis whether participants classified as AD positive according to a CSF p-tau/A $\beta$ 42 ratio threshold of .038 (Van Hulle et al., 2021), we also ran *post hoc* regressions within the AD positive cohort only (see Supplementary information for full results). These extra analyses were consistent with what reported above. When also evaluating inclusion probabilities, Rr is the best predictor for all outcomes, except for A $\beta$ 42, A $\beta$ 42/A $\beta$ 40, and NFL, as with the full sample. To note, as these analyses were based on a smaller sample, the findings should be interpreted with increased caution.

Unlike Bruno, Mueller, et al. (2021), who showed that Pr was predictive of amyloid load, as measured via Pittsburgh compound-B (PiB) positron emission tomography (PET), we did not observe an association between Pr and A $\beta$ 42 or the A $\beta$ 42/A $\beta$ 40 ratio. In this regards, we wish to make two observations. First, while the data were drawn in both cases from studies based at the University of Wisconsin – Madison, the actual samples were different: in the Bruno, Mueller, et al. (2021) paper, participants came from the Wisconsin Registry for Alzheimer's Prevention; these volunteers are generally younger and there is a higher proportion of cognitively

unimpaired individuals, compared to WADRC. Second, in the Bruno, Mueller, et al. (2021) study, the outcome was discrete, based on relevant PiB PET cut-points, whereas in the present study we examined continuous CSF levels as outcomes: we have noted that CSF and PiB PET markers will sometimes show differential levels of sensitivity to different cognitive (process) scores, and we plan on pursuing this observation further in the near future.

Limitations of this research should be noted. First of all, the sample sizes for MCI-AD and AD are significantly smaller than for the cognitively unimpaired participants. Sample sizes were dictated by availability, and future research with larger groups of individuals with cognitive impairment, possibly also including dementia pathologies other than AD, would be ideal to further these research questions. A second limitation is that the present sample nearly exclusively comprised individuals that identified as white Caucasians. While this may be positive methodologically, as possible confounding variables related to race are limited, many studies have highlighted the importance of including a wider spectrum of ethnicities and backgrounds in AD research (Manly et al., 2021; Morris et al., 2019). As far as we are aware, at least with regards to published works, Rr to date has only been tested primarily in white Caucasian populations – future research should consider examining whether the same patterns observed here would also extend to a more heterogeneous sample.

To conclude, this study showed that Rr, the ratio between immediate and delayed performance scores at the recency position is applicable to story recall, and sensitive to CSF levels of A $\beta$ 40, p-tau, t-tau, NFL, Ng and a-syn. Higher Rr scores, showing disproportionate loss of recency recall from immediate to delayed testing, were associated with worse biomarkers profiles, when controlling for age, diagnosis and APOE risk – and that the best predictors of biomarkers outcomes tended to be Rr combined with lower levels of immediate or delayed LMT performance. Therefore, we suggest the following: 1) Rr is a worthwhile measure to add to the clinician's battery (see also Egeland, 2021) when evaluating individuals suspected to be on a trajectory towards neurodegeneration; and 2) serial position values should be included in databases examining AD and other types of dementia. Future research should consider also comparing the relative predictive power of Rr when derived from word lists versus story recall; and examine whether the neurocognitive basis of Rr is different in word lists and story recall tasks.

---

### Credit author statement

Davide Bruno: Conceptualisation; formal analysis; funding acquisition; Writing – original draft. Ainara Jauregi Zinkunegi: formal analysis; Writing – original draft. Gwendlyn Kollmorgen: Methodology. Ivonne Suridjan: Methodology. Norbert Wild: Methodology. Cynthia Carlsson: Project administration. Barbara Bendlin: Writing – review & editing. Ozioma Okonkwo: Writing – review & editing. Nathaniel Chin: Writing – review & editing. Bruce P. Hermann: Writing – review & editing. Sanjay Asthana: Project administration. Henrik Zetterberg: Methodology; Writing – review & editing. Kaj Blennow: Methodology; Writing – review & editing. Rebecca

Langhough: Formal analysis; Writing – review & editing. Sterling C. Johnson: Project administration. Kimberly D. Mueller: Funding acquisition; Writing – review & editing.

### Author notes

The ethical regulations that govern the WADRC prevent unrestricted public archiving of anonymised study data. Data can be requested from the WADRC Executive Committee at: <https://www.adrc.wisc.edu/apply-resources>. Data will be released to internal and external investigators following confirmation of IRB approval together with an evaluation by the WADRC of scientific merit and resource availability.

This secondary analysis of WADRC data was funded by a NIH-NIA (R01 144 AAI8612) grant to KDM, in which DB and RL are co-investigators.

### Conflicts of interests and disclosure statement

IS is a full-time employee and shareholder of Roche Diagnostics International Ltd., GK is a full-time employee of Roche Diagnostics GmbH, and NW is a full-time employee of Roche Diagnostics GmbH. COBAS, COBAS E and ELECSYS are trademarks of Roche. The Elecsys®  $\beta$ -Amyloid (1–42) CSF assay, the Elecsys® Phospho-Tau (181P) CSF assay and the Elecsys® Total-Tau CSF assay are not approved for clinical use in the US. The NeuroToolKit robust prototype assays are for investigational purposes and are not approved for clinical use.

HZ is a Wallenberg Scholar supported by grants from the Swedish Research Council (#2018-02532), the European Research Council (#681712 and #101053962), Swedish State Support for Clinical Research (#ALFGBG-71320), the Alzheimer Drug Discovery Foundation (ADDF), USA (#201809-2016862), the AD Strategic Fund and the Alzheimer's Association (#ADSF-21-831376-C, #ADSF-21-831381-C and #ADSF-21-831377-C), the Olav Thon Foundation, the Erling-Persson Family Foundation, Stiftelsen för Gamla Tjänarinnor, Hjärfonden, Sweden (#FO2019-0228), the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860197 (MIR-IADE), the European Union Joint Programme – Neurodegenerative Disease Research (JPND2021-00694), and the UK Dementia Research Institute at UCL (UKDRI-1003).

KB is supported by the Swedish Research Council (#2017-00915), the Alzheimer Drug Discovery Foundation (ADDF) USA (#RDAPB-201809-2016615), the Swedish Alzheimer Foundation (#AF-930351, #AF-939721 and #AF-968270), Hjärfonden, Sweden (#FO2017-0243 and #ALZ2022-0006), the Swedish state under the agreement between the Swedish government and the County Councils, the ALF-agreement (#ALFGBG-715986 and #ALFGBG-965240), the European Union Joint Program for Neurodegenerative Disorders (JPND2019-466-236), the National Institute of Health (NIH) USA (grant #1R01AG068398-01), and the Alzheimer's Association 2021 Zenith Award (ZEN-21-848495).

No other author reports any conflicts of interests or disclosures.

### Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cortex.2022.12.004>.

### REFERENCES

- Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., Snyder, P. J., Carrillo, M. C., Thies, B., & Phelps, C. H. (2011). The diagnosis of mild cognitive impairment 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), 270–279.
- Alzheimer's Association. (2017). 2017. Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, 13, 325–373.
- Bock, J. R., Russell, J., Hara, J., & Fortier, D. (2021). Optimizing cognitive assessment outcome measures for Alzheimer's disease by matching wordlist memory test features to scoring methodology. *Frontiers in Digital Health*, 3.
- Bruno, D., Gleason, C. E., Kosciak, R. L., Pomara, N., Zetterberg, H., Blennow, K., & Johnson, S. C. (2019). The recency ratio is related to CSF amyloid beta 1-42 levels in MCI-AD. *International Journal of Geriatric Psychiatry*, 34(3), 415–419.
- Bruno, D., Jauregi, A., Pomara, N., Zetterberg, H., Blennow, K., Kosciak, R., Carlsson, C., Bendlin, B., Okonkwo, O., Hermann, B., Johnson, S., & Mueller, K. D. (2022). Cross-sectional associations of CSF tau levels with Rey's AVLT: A recency ratio study. *Neuropsychology*. Ahead of print <https://pubmed.ncbi.nlm.nih.gov/35604714/>.
- Bruno, D., Kosciak, R. L., Woodard, J. L., Pomara, N., & Johnson, S. C. (2018). The recency ratio as predictor of early MCI. *International Psychogeriatrics*, 30(12), 1883–1888.
- Bruno, D., Mueller, K. D., Betthausen, T., Chin, N., Engelman, C. D., Christian, B., Kosciak, R. L., & Johnson, S. C. (2021b). Serial position effects in the Logical Memory Test: Loss of primacy predicts amyloid positivity. *Journal of Neuropsychology*, 15(3), 448–461.
- Bruno, D., Reichert Plaska, C., Clark, D. P., Zetterberg, H., Blennow, K., Verbeek, M. M., & Pomara, N. (2021a). CSF  $\alpha$ -synuclein correlates with CSF neurogranin in late-life depression. *International Journal of Neuroscience*, 131(4), 357–361.
- Bruno, D., Reichert, C., & Pomara, N. (2016). The recency ratio as an index of cognitive performance and decline in elderly individuals. *Journal of Clinical and Experimental Neuropsychology*, 38(9), 967–973.
- Campbell, M. R., Ashrafzadeh-Kian, S., Petersen, R. C., Mielke, M. M., Syrjanen, J. A., van Harten, A. C., Lowe, V. J., Jack, C. R., Jr., Bornhorst, J. A., & Algeciras-Schimnich, A. (2021). P-tau/A $\beta$ 42 and A $\beta$ 42/40 ratios in CSF are equally predictive of amyloid PET status. *Alzheimer's & Dementia*, 13(1), Article e12190. <https://doi.org/10.1002/dad2.12190>. Published 2021 May 18.
- Carlesimo, G. A., Sabbadini, M., Fadda, L., & Caltagirone, C. (1995). Different components in word-list forgetting of pure amnesics, degenerative demented and healthy subjects. *Cortex*, 31(4), 735–745.
- Darst, B. F., Kosciak, R. L., Racine, A. M., Oh, J. M., Krause, R. A., Carlsson, C. M., Zetterberg, H., Blennow, K., Christian, B. T., Bendlin, B. B., Okonkwo, O. C., Hogan, K. J., Hermann, B. P., Sager, M. A., Asthana, S., Johnson, S. C., & Engelman, C. D. (2017). Pathway-specific polygenic risk scores as predictors of amyloid- $\beta$  deposition and cognitive function in a sample at



- increased risk for Alzheimer's disease. *Journal of Alzheimer's Disease*, 55(2), 473–484.
- Egeland, J. (2021). Following HN over 21 years: Recency change and reduced retention predict later impairment in memory, and recency ratio may combine both effects. *Neurocase*, 27(2), 147–154.
- Foldi, N. S., Brickman, A. M., Schaefer, L. A., & Knutelska, M. E. (2003). Distinct serial position profiles and neuropsychological measures differentiate late life depression from normal aging and Alzheimer's disease. *Psychiatry Research*, 120(1), 71–84.
- Hall, S., & Bornstein, R. A. (1991). Serial-position effects in paragraph recall following mild closed-head injury. *Perceptual and Motor Skills*, 72(3, suppl.), 1295–1298.
- Jauregi, A., Hermann, B. P., Betthausen, T. J., Langhough Koscik, R., Johnson, S. C., Mueller, K. D., & Bruno, D. (2022). Proximal recency ratio in story recall predicts left hippocampal volume in older adults with and without cognitive impairment. In *Poster presented at the spring British Neuropsychological Society meeting*, London, UK.
- Li, X., Li, T. Q., Andreasen, N., Wiberg, M. K., Westman, E., & Wahlund, L. O. (2013). Ratio of A $\beta$ 42/P-tau<sub>181p</sub> in CSF is associated with aberrant default mode network in AD. *Scientific Reports*, 3, 1339. <https://doi.org/10.1038/srep01339>
- Maass, A., Berron, D., Harrison, T. M., Adams, J. N., La Joie, R., Baker, S., Mellinger, T., Bell, R. K., Swinnerton, K., Inglis, B., Rabinovici, G. D., Düzel, E., & Jagust, W. J. (2019). Alzheimer's pathology targets distinct memory networks in the ageing brain. *Brain*, 142(8), 2492–2509.
- Manly, J. J., Avila, J. F., Vonk, J. M., Rentería, M. A., Turney, I. C., Lao, P. J., Seblova, D., Martínez, M. N., Gutierrez, J., Mayeux, R., Schupf, N., & Brickman, A. M. (2021). Racial differences in cognitive resilience to parental history of AD and cognitive impairment: The Offspring study. *Alzheimer's & Dementia*, 17, Article e051874.
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Jr., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., Mohs, R. C., Morris, J. C., Rossor, M. N., Scheltens, P., Carrillo, M. C., Thies, B., Weintraub, S., & Phelps, C. H. (2011). 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), 263–269.
- Morris, J. C., Schindler, S. E., McCue, L. M., Moulder, K. L., Benzinger, T. L. S., Cruchaga, C., Fagan, A. M., Grant, E., Gordon, B. A., Holtzman, D. M., & Xiong, C. (2019). Assessment of racial disparities in biomarkers for Alzheimer disease. *JAMA Neurology*, 76(3), 264–273.
- Mueller, K. D., Du, L., Bruno, D., Betthausen, T., Christian, B., Johnson, S., ... Koscik, R. L. (2022). Item-level story recall predictors of amyloid-beta in late middle-aged adults at increased risk for Alzheimer's disease. *Frontiers in Psychology*, 13, 908651. <https://doi.org/10.3389/fpsyg.2022.908651>. PMID: 35832924; PMCID: PMC9271832.
- Murdock, B. B., Jr. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, 64(5), 482.
- Selnes, P., Stav, A. L., Johansen, K. K., Bjørnerud, A., Coello, C., Auning, E., Kalheim, L., Almdahl, I. S., Hessen, E., Zetterberg, H., Blennow, K., & Fladby, T. (2017). Impaired synaptic function is linked to cognition in Parkinson's disease. *Annals of Clinical and Translational Neurology*, 4(10), 700–713.
- Shim, K. H., Kang, M. J., Suh, J. W., Pyun, J. M., Ryoo, N., Park, Y. H., Youn, Y. C., Jang, J. W., Jeong, J. H., Park, K. W., & Choi, S. H. (2020). CSF total tau/ $\alpha$ -synuclein ratio improved the diagnostic performance for Alzheimer's disease as an indicator of tau phosphorylation. *Alzheimer's Research & Therapy*, 12(1), 1–12.
- Teipel, S. J., Dyrba, M., Ballarini, T., Brosseron, F., Bruno, D., Buerger, K., Cosma, N.-C., Dechent, P., Dobisch, L., Düzel, E., Ewers, M., Fließbach, K., Haynes, J. D., Janowitz, D., Kilimann, I., Laske, C. D., Maier, F., Metzger, C., & Munk, M. H. (2021). Association of cholinergic basal forebrain volume and functional connectivity with markers of inflammatory response in the Alzheimer's disease spectrum. *Journal of Alzheimer's Disease*, 1–16 (Preprint).
- Tennant, V. R., Harrison, T. M., Adams, J. N., La Joie, R., Winer, J. R., & Jagust, W. J. (2021). Fusiform gyrus phospho-tau is associated with failure of proper name retrieval in aging. *Annals of Neurology*, 90(6), 988–993.
- Turchetta, C. S., De Simone, M. S., Perri, R., Fadda, L., Caruso, G., De Tollis, M., Caltagirone, C., & Carlesimo, G. A. (2020). Forgetting rates on the recency portion of a word list predict conversion from mild cognitive impairment to Alzheimer's disease. *Journal of Alzheimer's Disease*, 73(4), 1295–1304.
- Turchetta, C. S., Perri, R., Fadda, L., Caruso, G., De Simone, M. S., Caltagirone, C., & Carlesimo, G. A. (2018). Forgetting rate on the recency portion of a word list differentiates mild to moderate Alzheimer's disease from other forms of dementia. *Journal of Alzheimer's Disease*, 66(2), 461–470.
- Van Hulle, C., Jonaitis, E. M., Betthausen, T. J., Batrla, R., Wild, N., Kollmorgen, G., Andreasson, U., Okonkwo, O., Bendlin, B. B., Asthana, S., Carlsson, C. M., Johnson, S. C., Zetterberg, H., & Blennow, K. (2021). An examination of a novel multipanel of CSF biomarkers in the Alzheimer's disease clinical and pathological continuum. *Alzheimer's & Dementia*, 17(3), 431–445.
- Wechsler, D. (1987). *WMS-R: Wechsler memory scale-revised*. Psychological Corporation.
- Wellington, H., Paterson, R. W., Portelius, E., Törnqvist, U., Magdalinou, N., Fox, N. C., Blennow, K., Schott, J. M., & Zetterberg, H. (2016). Increased CSF neurogranin concentration is specific to Alzheimer disease. *Neurology*, 86(9), 829–835. <https://doi.org/10.1212/WNL.0000000000002423>
- Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology*, 55, 235–269.
- Wixted, J. T., & Cai, D. J. (2013). Memory consolidation. In *Oxf. Handb. Cogn. Neurosci* (Vol. 1, pp. 436–456).
- Zetterberg, H., & Bendlin, B. B. (2021). Biomarkers for Alzheimer's disease—preparing for a new era of disease-modifying therapies. *Molecular Psychiatry*, 26(1), 296–308.