JCI The Journal of Clinical Investigation

Wet or dry: translatable “water mazes” for mice and humans

Kerin K. Higa, … , Jared W. Young, Mark A. Geyer

J Clin Invest. 2016;126(2):477-479. https://doi.org/10.1172/JCI86071.

Commentary

Although the cognitive and biological characteristics of Alzheimer’s disease (AD) are well known and mouse models of AD are available, current treatments for AD-related cognitive deficits have quite limited efficacy. The development of tasks with cross-species validity may enable better prediction of the efficacy of potential new treatments. In this issue of the *JCI*, Possin et al. present a virtual version of the Morris water maze (a common test of spatial learning and memory for rodents) that is designed for use with humans. The authors tested a mouse model of AD (transgenic mice expressing human amyloid precursor protein [hAPP]) and patients in the earlier mild cognitive impairment (MCI) stage of AD in their respective versions of the maze. Using novel statistical methods, they detected similar deficits across species, providing support for the hAPP model and use of the virtual water maze. Importantly, this work enabled recommendations for appropriate sample sizes when developing potential therapeutics for AD.

Wet or dry: translatable "water mazes" for mice and humans

Kerin K. Higa,[1] Jared W. Young,[1,2] and Mark A. Geyer[1,2]

[1]Department of Psychiatry, University of California San Diego (UCSD), La Jolla, California, USA. [2]Research Service, VA San Diego Healthcare System, San Diego, California, USA.

Although the cognitive and biological characteristics of Alzheimer's disease (AD) are well known and mouse models of AD are available, current treatments for AD-related cognitive deficits have quite limited efficacy. The development of tasks with cross-species validity may enable better prediction of the efficacy of potential new treatments. In this issue of the *JCI*, Possin et al. present a virtual version of the Morris water maze (a common test of spatial learning and memory for rodents) that is designed for use with humans. The authors tested a mouse model of AD (transgenic mice expressing human amyloid precursor protein [hAPP]) and patients in the earlier mild cognitive impairment (MCI) stage of AD in their respective versions of the maze. Using novel statistical methods, they detected similar deficits across species, providing support for the hAPP model and use of the virtual water maze. Importantly, this work enabled recommendations for appropriate sample sizes when developing potential therapeutics for AD.

A disconnect between animal models and human AD patients

Alzheimer's disease (AD) is a devastating illness and the leading cause of dementia in older adults. As with other neurological disorders, the study of AD relies on animal models in conjunction with human studies to explore the pathogenesis of the disease and to develop treatments to slow or prevent disease progression. Thus far, treatments with agents such as acetylcholinesterase (AChE) inhibitors to alleviate the cognitive deficits associated with AD produce only modest cognitive improvements that are neither permanent nor preventative of further decline (1, 2). To better predict the efficacy of therapeutic strategies, it is necessary to develop translatable paradigms that reliably model cognitive impairment (and improvement) across species, as discussed for other diseases with cognitive deficits such as schizophrenia (3, 4).

The pathogenesis and biological signatures of AD are well established, and several mouse lines have been developed to mimic β amyloid plaque and tau neurofibrillary tangle formation (5, 6). These models often exhibit cognitive deficits similar to those observed in patients at various stages of AD. Directly translatable comparisons are limited, however, because rodents are assessed with ethologically relevant tasks, while patients are tested using pen and paper.

For rodents, the most commonly used test of spatial learning and memory is the Morris water maze (MWM). There are several versions of the MWM designed to assess various elements of spatial learning and memory, and several mouse models of AD exhibit impaired performance in the MWM. Classic human tests rely more on memory tests that are based on short stories or visual image reproductions (5). Recently, the MWM has been reverse translated for use in humans, providing evidence that patients with AD are impaired in two-dimensional, real space, and virtual reality versions of the task (7–10).

Although performance in the MWM of both mice and humans relies on hippocampal integrity (7, 11), especially the right hippocampus for AD patients (12), the extent to which rodent MWM performance can predict and be compared with human performance was heretofore unknown. In this issue, Possin et al. (13) assessed the translatability of MWM findings by using comparable mouse and human MWM paradigms and measures (Table 1 and Figure 1), enabling direct comparisons across species.

Performance comparison of species-specific MWM paradigms

Both mouse and human versions had three test stages: visible target training, hidden target learning, and a probe test. The human test took place in a virtual circular field that was navigable with a steering wheel. This methodology enabled first-person exploration of the space to find a reward. Mice, however, had to navigate through water to find first the visible, then hidden escape platforms. The performance measures analyzed for the visible and hidden target stages were distance (mean proximity to the target), latency (percentage of time spent in the target quadrant), and cumulative search error (CSE) (the cumulative distance from the platform, collected in 1-s averages) (see ref. 14). During the probe test, the performance measures were mean proximity to the target location and percentage of time spent in the target quadrant. Hence, the outcome variables were consistent across species.

To analyze these outcome variables, Possin and colleagues developed and assessed the use of "rank summary scores"

▶ Related Article: p. 779

Conflict of interest: In the past three years, M.A. Geyer has received consulting compensation from Abbott, Dart NeuroScience LLC, Lundbeck Ltd., Neurocrine Biosciences Inc., Omeros Corp., Otsuka, and Sunovion Pharmaceuticals Inc. and holds an equity interest in San Diego Instruments. J.W. Young's work has been funded by Lundbeck Ltd. and Omeros, and he has received consulting compensation from Amgen and Sunovion.
Reference information: *J Clin Invest*. 2016;126(2):477–479. doi:10.1172/JCI86071.

Table 1. Timeline for MWM experiments with AD model mice and humans with MCI-AD

	Visible target training	Hidden target learning	Probe test
Mouse	2 trials × 2 sessions × 3 days. Max. time: 60 s	2 trials × 6 days. Max. time: 90 s	18–20 h after last session. Duration: 90 s
Human	4 trials. Max. time: 5 s	10 trials. Max. time: 120 s	40 min after last session. Duration: 90 s

Max., maximum.

to compare group performance measures during the learning stages. Traditionally, a repeated-measures ANOVA is used to analyze performance measures across trials; however, this does not account for nonlinear learning curves, trial-to-trial variability in performance, or aborted trials, all of which commonly lead to unreliable findings. The use of rank summary scores enabled comparison between groups using a simple two-sample Student's *t* test. Rank summary scores were calculated by replacing raw scores with quantile scores (each subject's ranking, divided by the total number of subjects), which were then averaged across trials. Theoretically, this method should reduce the influence of outliers and eliminate the effects of nonlinear learning, nonconstant trial-by-trial variance, and right censoring due to trials aborted when the target was not found. The use of rank summary scores also enabled direct cross-species comparisons.

Possin et al. assessed the comparability of the mouse and human MWM protocols by testing the performance of a mouse model of AD — transgenic mice expressing human amyloid precursor protein (hAPP of the J20 line) — as well as patients diagnosed with mild cognitive impairment (MCI) who were likely to develop AD dementia (MCI-AD). This line of hAPP mice develops cognitive impairments as early as 1 to 2 months of age, with spatial learning and memory deficits apparent by 6 months, and may be a good model of AD before the onset of dementia (6, 15, 16). Therefore, the hAPP mice and MCI-AD patient group might develop similar cognitive deficits (5, 17). Indeed, both MCI-AD patients and hAPP mice showed poorer hidden target and probe trial performance. In contrast to MCI-AD patients, however, hAPP mice exhibited deficits in visible platform testing, as reported inconsistently before (18–20). Possin and colleagues attributed the learning deficit to difficulty in relinquishing a thigmotactic response (swimming around the outside of the pool), as the water tank is stressful for the mice. The difference in motivational drive between the species and the use of verbal instructions for humans but not mice could underlie this difference in visible platform learning across species.

Importantly, key deficits in hidden target learning were similar between MCI-AD patients and hAPP mice. Both groups had greater average rank summary scores for distance, latency, and CSE (learning scores). During the probe test, both groups had greater mean proximity to the platform and spent a smaller percentage of time in the target quadrant (memory score). Using these data, Possin et al. also made recommendations for the group sizes necessary to detect treatment effects in both mouse and human test groups. The simple methods suggested by Possin et al. might enable faster and more reliable screening for AD treatments that can be performed across species, theoretically enhancing the translation of procognitive findings in rodents to the clinical patient population. In addition, these analyses could benefit other studies examining nonlinear learning tasks.

Conclusions and future directions

These results demonstrate the value of consistently using behavioral tests and parameters that can be conducted in humans and rodents and provide strong support for the hAPP mouse model of AD. Several key differences in task design should be noted, however. As already intimated, the difference in motivational drive could underlie differences in pharmacological effects (Table 1 and Figure 1). Karl et al. (17) addressed this issue by using the cheeseboard task, a "dry version" of the MWM, in which the mice had to remember the location of a food reward on a circular platform. There were no differences between hAPP mice and controls in the training phase (though using rank summary scores might reveal a difference), but the hAPP mice spent significantly less time in the target quadrant during the probe test than did the controls. Use of this dry version with some adjustments to the human task would help to equate motivational drive, albeit with the added confounding factor of food deprivation for the mice. Alternatively, the human MWM could be altered so that the subjects

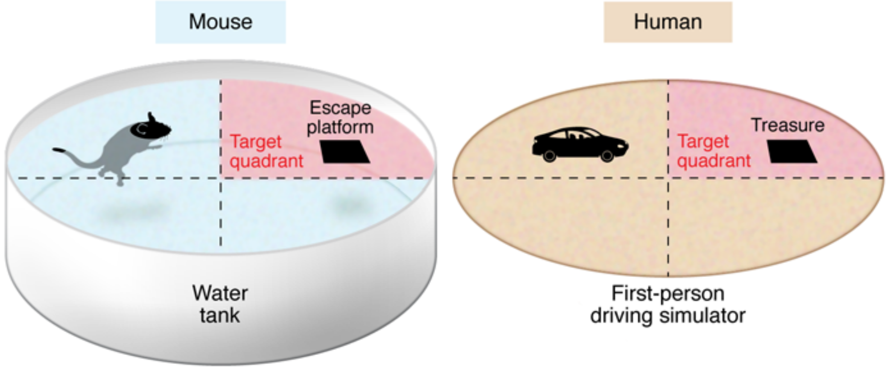

Figure 1. Schematic of hidden target learning arenas for mouse and human tasks. The mouse tasks were performed over several days, while all stages of the human tasks were performed on a single day. Mice were trained to swim to a hidden escape platform. Humans were trained to drive to a hidden treasure chest using a driving simulator that was situated in front of a computer monitor. For hidden target learning and probe tests, each arena was surrounded by consistent landmarks for spatial reference (not pictured).

must similarly escape from (virtual) danger, such as a tidal wave. In addition to motivational differences, the human and mouse versions of the test differ in their timelines (Table 1 and Figure 1). Mice were tested over several days, probably requiring greater hippocampal processing due to long-term memory consolidation, whereas the humans were tested on a single day, requiring only short-term memory. These limitations did not impact the similar deficits seen across these species, however, indicating that such translational comparisons are possible.

Although AChE inhibitors, which are currently approved as treatments for AD, produce only moderate benefits in patients, it would be informative to test both mice and patients in the MWM after AChE inhibitor treatment, as this would enable assessment of the predictive validity of both the hAPP mouse model and the MWM paradigm. The inclusion of these tasks in ongoing studies for newer treatments, such as anti-amyloid and anti-tau immunotherapies, would also be informative (1). Additionally, certain genetic crosses have been shown to attenuate spatial memory deficits in this hAPP mouse line (19–22). Although such manipulations cannot be achieved in patients, they might be useful to determine the sensitivity of each performance measure to previously established genetic rescues. Future studies should investigate other AD models and treatments using these standardized techniques, and many past studies could be reanalyzed accordingly.

Further investigation of the underlying mechanisms that cause deficits in MWM performance by patients and animal models is warranted. For example, the poor performance of the mice on the visible target test could have been due to poor vision, although previous studies do not indicate vision problems in these animals. Additionally, an increased tendency for thigmotactic swimming (22) might indicate anxiety, helplessness, or cognitive inflexibility. If these mice do indeed have impaired flexibility, perhaps their poor performance on the hidden maze task is due to a continued search for the visible target. Furthermore, considering that the cognitive deficits in AD extend beyond such learning and memory, including attention and executive functioning deficits, future studies could test these mice and patients on available batteries of cross-species tasks, as has been discussed for other diseases (3, 23). The approach highlighted here by Possin et al. provides an excellent first step toward directly translational research that can increase the likelihood of procognitive treatments that bridge the bench-to-bedside gap.

Acknowledgments

This work was supported by NIH grants R21-MH081037, R21-MH101579, R01-MH104344, and F31-MH109218 and by the VA's VISN 22 Mental Illness Research, Education, and Clinical Center.

Address correspondence to: Jared W. Young, Department of Psychiatry, University of California San Diego, 9500 Gilman Drive MC 0804, La Jolla, California 92093-0804, USA. Phone: 619.543.3582; E-mail: jaredyoung@ucsd.edu.

1. Rafii MS, Aisen PS. Advances in Alzheimer's disease drug development. *BMC Med*. 2015;13(1):62.
2. McGleenon BM, Dynan KB, Passmore AP. Acetylcholinesterase inhibitors in Alzheimer's disease. *Br J Clin Pharmacol*. 1999;48(4):471–480.
3. Young J, Geyer M. Developing treatments for cognitive deficits in schizophrenia: the challenge of translation. *J Psychopharmacol*. 2015;29(2):178–196.
4. Moore H, Geyer MA, Carter CS, Barch DM. Harnessing cognitive neuroscience to develop new treatments for improving cognition in schizophrenia: CNTRICS selected cognitive paradigms for animal models. *Neurosci Biobehav Rev*. 2013;37(9 pt B):2087–2091.
5. Webster SJ, Bachstetter AD, Nelson PT, Schmitt FA, Van Eldik LJ. Using mice to model Alzheimer's dementia: an overview of the clinical disease and the preclinical behavioral changes in 10 mouse models. *Front Genet*. 2014;5:88.
6. Ashe KH, Zahs KR. Probing the biology of Alzheimer's disease in mice. *Neuron*. 2010;66(5):631–645.
7. Burgess N, Trinkler I, King J, Kennedy A, Cipolotti L. Impaired allocentric spatial memory underlying topographical disorientation. *Rev Neurosci*. 2006;17(1–2):239–251.
8. deIpolyi AR, Rankin KP, Mucke L, Miller BL, Gorno-Tempini ML. Spatial cognition and the human navigation network in AD and MCI. *Neurology*. 2007;69(10):986–997.
9. Boccia M, Nemmi F, Guariglia C. Neuropsychology of environmental navigation in humans: review and meta-analysis of FMRI studies in healthy participants. *Neuropsychol Rev*. 2014;24(4):236–251.
10. Laczó J et al. Spatial navigation testing discriminates two types of amnestic mild cognitive impairment. *Behav Brain Res*. 2009;202(2):252–259.
11. Vlček K, Laczó J. Neural correlates of spatial navigation changes in mild cognitive impairment and Alzheimer's disease. *Front Behav Neurosci*. 2014;8:89.
12. Nedelska Z, et al. Spatial navigation impairment is proportional to right hippocampal volume. *Proc Natl Acad Sci U S A*. 2012;109(7):2590–2594.
13. Possin KL, et al. Cross-species translation of the Morris maze for Alzheimer's disease. *J Clin Invest*. 2016;126(2):779–783.
14. Gallagher M, Burwell R, Burchinal M. Severity of spatial learning impairment in aging: development of a learning index for performance in the Morris water maze. *Behav Neurosci*. 1993;107(4):618–626.
15. Palop JJ, et al. Neuronal depletion of calcium-dependent proteins in the dentate gyrus is tightly linked to Alzheimer's disease-related cognitive deficits. *Proc Natl Acad Sci U S A*. 2003;100(16):9572–9577.
16. Harris JA, et al. Many neuronal and behavioral impairments in transgenic mouse models of Alzheimer's disease are independent of caspase cleavage of the amyloid precursor protein. *J Neurosci*. 2010;30(1):372–381.
17. Mucke L, et al. High-level neuronal expression of aβ 1-42 in wild-type human amyloid protein precursor transgenic mice: synaptotoxicity without plaque formation. *J Neurosci*. 2000;20(11):4050–4058.
18. Karl T, Bhatia S, Cheng D, Kim WS, Garner B. Cognitive phenotyping of amyloid precursor protein transgenic J20 mice. *Behav Brain Res*. 2012;228(2):392–397.
19. Sanchez-Mejia RO, et al. Phospholipase A2 reduction ameliorates cognitive deficits in a mouse model of Alzheimer's disease. *Nat Neurosci*. 2008;11(11):1311–1318.
20. Roberson ED, et al. Reducing endogenous tau ameliorates amyloid beta-induced deficits in an Alzheimer's disease mouse model. *Science*. 2007;316(5825):750–754.
21. Poirier R, et al. Neuronal neprilysin overexpression is associated with attenuation of Aβ-related spatial memory deficit. *Neurobiol Dis*. 2006;24(3):475–483.
22. Galvan V, et al. Long-term prevention of Alzheimer's disease-like behavioral deficits in PDAPP mice carrying a mutation in Asp664. *Behav Brain Res*. 2008;191(2):246–255.
23. Keeler JF, Robbins TW. Translating cognition from animals to humans. *Biochem Pharmacol*. 2011;81(12):1356–1366.