Published in final edited form as:
*Behav Brain Res*. 2010 September 1; 212(1): 115–120. doi:10.1016/j.bbr.2010.03.037.

# Effect Size of Reference Memory Deficits in the Morris Water Maze in Tg2576 Mice

**Miranda N. Reed**[a,b], **Peng Liu**[a,b], **Linda A. Kotilinek**[a,b], and **Karen H. Ashe**[a,b,c,*]

[a] N. Bud Grossman Center, University of Minnesota, Minneapolis, MN, 55455

[b] Departments of Neurology, University of Minnesota, Minneapolis, MN, 55455

[c] Geriatric Research, Education and Clinical Center, Veterans Affairs Medical Center, Minneapolis, MN 55417

## Abstract

The most widely used mouse model of Alzheimer's disease is the Tg2576 ($APP_{SWE}$) model. While general agreement about their neuropathology prevails, disparate results concerning cognitive changes have been reported. To resolve this controversy, we combined Morris water maze data collected over >10 years to determine the extent of memory impairment. $APP_{SWE}$ mice exhibited an age-dependent decline in memory, but the effect size was small when compared to non-transgenic littermates. Larger effect sizes were achieved when comparing $APP_{SWE}$ and Tg5469 ($APP_{WT}$) mice.

### Keywords

Tg2576; Alzheimer's disease; sample size; effect size; Morris Water Maze

Alzheimer's disease (AD) is characterized by an age-related impairment in learning and memory, neuronal loss, gliosis, neuritic changes, amyloid deposition, and abnormal tau phosphorylation and aggregation [1–3]. Animal models of AD should display both the pathological changes observed in AD, as well as changes in memory function that worsen in an age-dependent manner. The latter is particularly important since the primary risk factor for sporadic AD is age.

Over 15 years ago, the Tg2576 mouse (herein referred to as $APP_{SWE}$) was developed as an animal model of AD, and has since been used in over 607 articles pertaining to the pathogenesis and treatment of AD. This model over-expresses the 695 amino acid human isoform of the

*Corresponding Author at: Department of Neurology, MMC 295, Center for Memory Research & Care, Department of Neurology, University of Minnesota, 2101 6th St SE, Minneapolis, MN 55455, USA, hsiao005@umn.edu, Tel: 612-626-0652, Fax: 612-626-2639.

**Authors' Information**

Miranda N. Reed, Center for Memory Research & Care, Department of Neurology, University of Minnesota, 2101 6th St SE, Minneapolis, MN 55455, USA

Peng Liu, Center for Memory Research & Care, Department of Neurology, University of Minnesota, 2101 6th St SE, Minneapolis, MN 55455, USA

Linda A. Kotilinek, Center for Memory Research & Care, Department of Neurology, University of Minnesota, 2101 6th St SE, Minneapolis, MN 55455, USA

Karen H. Ashe, Center for Memory Research & Care, Department of Neurology, University of Minnesota, 2101 6th St SE, Minneapolis, MN 55455, USA

amyloid precursor protein ($APP_{695}$) with the "Swedish" mutation, resulting in the overproduction of the amyloid-beta (Aβ) peptide [4]. The $APP_{SWE}$ mouse model recapitulates some of the hallmarks of AD, including neuritic changes, inflammation, amyloid deposition and plaques [5–7]. Although most studies have found memory deficits in the $APP_{SWE}$ mice, discrepancies exist as to when the onset of these deficits first occurs, with reports ranging from 3 months to as late as 15 months, as well as ages in-between [4,8–11]. Disparate results occur within laboratories as well [4,8,12–14]. Some investigators have failed to find deficits altogether [15]. Thus, it has been difficult to discern if an age-related impairment exists in the $APP_{SWE}$ model.

Several possible reasons for these discrepancies exist, including sensitivity differences in the cognitive tests used. The current paper, however, focuses on two other possibilities: (1) the relatively small effect size seen when comparing $APP_{SWE}$ mice to transgene negative (Tg Neg) mice using the Morris water maze and (2) the cognitive-enhancing effects of secreted APPα [sAPPα; 16,17] and the APP intracellular domain [AICD; 18], additional byproducts of APP cleavage.

In the $APP_{SWE}$ mouse model, sAPPα, AICD, and Aβ are over-expressed following proteolytic processing of APP. sAPPα has been shown to enhance long-term potentiation (LTP), modulate the induction of long-term depression (LTD) [16], and enhance memory performance in a variety of learning-tasks following intracerebroventricular injection [17]. Likewise, AICD facilitates memory and synaptic plasticity [18,19]. This stands in opposition to Aβ, which impairs memory [4,14,20] and synaptic function [9,21–24]. Consequently, the proper control for $APP_{SWE}$ mice is the Tg5469 ($APP_{WT}$) mouse line. We used previously described methods [20] to show that, with the exception of Aβ, the levels of APP and APP metabolites were similar between the two lines. Briefly, a four-step extraction protocol was used to generate four fractions (extracellular-enriched soluble (EC), intracellular-enriched soluble (IC), membrane-enriched (MB) and insoluble). 1 ug of protein from the EC and MB fractions were loaded onto gels to probe for sAPPα and APP, respectively, using 6E10 (Signet, 1:2500). 150 ug of protein from the MB was used to probe for AICD with an anti-APP C-terminal antibody 0443 (Millipore, 1:5000), and 100 ug of protein from the EC fraction was used to probe for Aβ with 6E10. Blots were stripped and reprobed for α-Tubulin using an anti-α-Tubulin antibody (Sigma, 1:200,000). OptiQuant was used for densitometric analysis, and bands were normalized to α–Tubulin loading controls for each sample. Using this method, we found that the $APP_{WT}$ mice over-express wild-type human APP at levels equivalent to mutant APP in $APP_{SWE}$ mice (Fig. 1A & 1E), have equivalently high levels of AICD (Fig. 1B & 1E) and sAPPα (Fig. 1C & 1E) but much lower levels of Aβ (Fig. 1D & 1E).

These are important issues because APP transgenic mice are often used to evaluate potential therapies for AD. Knowing when memory loss first appears as well as the appropriate sample size needed is essential for establishing experimental designs. In addition, an appropriate effect size is needed to ensure that the dynamic range for a particular cognitive task is large enough so that subtle treatment effects can be detected. Comparison of $APP_{SWE}$ to $APP_{WT}$ mice should increase the effect size of $APP_{SWE}$ mice by controlling for the beneficial effects of sAPPα and AICD overexpression, making this a more useful model for therapeutic testing. The current paper tests this possibility.

Here, we combined Morris water maze data collected over more than 10 years in our laboratory in order to determine if an age-dependent impairment in memory exists, and report effect size and sample size needed at various ages. We also report the effect sizes of $APP_{SWE}$ versus $APP_{WT}$ mouse models in an effort to distinguish the beneficial effects of sAPPα and AICD expression from the detrimental effects of Aβ.

Spatial reference learning and memory was tested using the Morris water maze in a cohort of 86 Tg5469 ($APP_{WT}$), 146 Tg2576 mice ($APP_{SWE}$), and 243 Tg negative littermates (Tg Neg) at various time points across their life span (Table 1). All transgenic mice used in this study were generated by breeding transgene positive male $APP_{WT}$ or $APP_{SWE}$ mice to female C57Bl6j/SJL F1 mice. The resultant mixed-background mice are F2-like in strain characteristics. To ensure that strain backgrounds in the $APP_{WT}$ and $APP_{SWE}$ lines were similar, behavioral scores for Tg Negs littermates from both lines were compared. There were no differences between the Tg Negs for any measure tested (*P*s>0.5).

Mice were grouped into four age ranges [14] based on previously established changes in soluble Aβ oligomers [20], detergent-insoluble Aβ ($Aβ_{insol}$) levels [25], and plaque pathology [4]: (1) very young mice, 4–5 months, after the appearance of Aβ trimers and hexamers but before the appearance of $Aβ_{insol}$ or plaques; (2) young mice, 6–11 months, after the appearance of Aβ*56, a 56-kDa Aβ oligomer, and during the initial appearance of $Aβ_{insol}$ and both amyloid plaques and punctate Aβ deposits; (3) middle-aged mice, 12–18 months, during a period in which soluble Aβ oligomers do not change but there is extensive deposition of plaques and $Aβ_{insol}$ levels are rising rapidly; and (4) old mice, 20–25 months, at a time when soluble Aβ levels are stable, $Aβ_{insol}$ is leveling off and amyloid loads are comparable to those in patients with AD (see Tables 1, 2). Mice were naïve at each time point tested.

We previously described in detail the Morris water maze procedure used here [14]. Briefly, at each age tested, mice received visible platform training for 3 days, eight trials per day, followed by hidden platform training for 9 days, four trials per day. Three probe trials of 60 s duration were performed at the beginning of the 4th, 7th, and 10th day of hidden platform training. The mean platform score (MPS) was used to assess retention of spatial reference memory and was calculated for each mouse by averaging time spent in the quadrant area for the three probes conducted. Although similar trends were observed using the platform crossing index (PCI), percent time is reported here because it is the most popular dependent measure reported in the Morris water maze literature [26], including our own search of the $APP_{SWE}$ literature.

All trials were monitored using a computerized tracking system (Noldus EthoVision 3.0; Noldus Information Technology, Wageningen, The Netherlands), and performance measures were extracted using Wintrack (Wolfer, et al. 2001). Statistical analysis consisted of t-tests, ANOVA and repeated-measures ANOVA (RMANOVA). *Post hoc* comparisons were performed using Bonferroni with *p* values of <0.05 considered significant.

Effect size and sample size needed was calculated for each probe and MPS using the software package Systat (2004). Cohen's *d* for t-tests was used to calculate effect sizes and was chosen for two reasons. First, this calculation is one of the most popular, allowing for easy comparison to other published studies. Second, Cohen's [27] classification of effect sizes into categories (.20 - small, .50 - medium, and .80 - large) makes evaluation of this experiment's effect-size results easy to compare to known benchmarks.

To calculate effect size, we computed the standardized mean difference (SMD) as the difference between the $APP_{SWE}$ mice and either the Tg Neg or the $APP_{WT}$ mice divided by the pooled standard deviation. In addition, we compared the Tg Neg mice to the $APP_{WT}$ mice.

$$d = M_1 - M_2 / \sigma_{pooled} \quad \text{(a)}$$

$$\sigma_{pooled} = \sqrt{[((n_1 - 1)s_1^2 + (n_2 - 1)s_2^2)/(n_1 + n_2)]} \quad \text{(b)}$$

Key to symbols:

***d*** = Cohen's ***d*** effect size

$\boldsymbol{M_1}$ = mean (average of Tg Neg or $APP_{WT}$)

$\boldsymbol{M_2}$ = mean (average of $APP_{SWE}$)

***s*** = standard deviation

***n*** = number of subjects

Effect sizes were computed as Cohen's *d* where a positive effect size represents better performance for the (1) Tg Neg mice when compared to either the $APP_{WT}$ or $APP_{SWE}$ mice or (2) $APP_{WT}$ when compared to the $APP_{SWE}$ mice (Tables 2 & 3). In order to estimate an appropriate sample size, the SMD was determined, the probability of a Type I error (α) and power were set at 0.05 and 0.80, respectively, and the alternative was specified as not equal.

To justify individual comparisons of transgene at each age range, we first examined the effect of age, gender, and transgene on probe scores to determine if there were any main effects or interactions among these variables. This analysis revealed a main effect of (1) age ($P$=0.004), performance declined as the mice aged, (2) transgene ($P$<0.0001), $APP_{WT}$ mice performed better than Tg Negs and $APP_{SWE}$ mice ($p$s<0.01) while Tg Negs were superior to $APP_{SWE}$ mice ($p$<0.01), (3) gender ($P$=0.0057), males performed better than females, and (4) an interaction between transgene and age ($P$<0.0001). To clarify the interaction between age and transgene, we examined the effects of transgene at each age range.

When compared with $APP_{WT}$ mice, spatial reference memory was significantly impaired in $APP_{SWE}$ mice at all ages tested after 4–5 months of age (Fig. 2). At 6–11 and 12–18 months of age, a medium to large effect size was observed for every measure when comparing $APP_{SWE}$ to $APP_{WT}$ mice, regardless of gender (Table 2 and 3). However, at 20–24 months of age the effect size was gender dependent; $APP_{WT}$ males performed better than $APP_{SWE}$ males at each probe, whereas $APP_{WT}$ females were more variable (Table 3). In contrast, comparison between Tg Neg and $APP_{SWE}$ mice revealed much smaller effect sizes, regardless of gender (Table 3), leading to a lack of statistical difference at 6–11 months of age (Fig. 2). In addition, these small effect sizes mandated the need for much larger sample sizes compared to the practical numbers needed for comparisons between $APP_{SWE}$ and $APP_{WT}$ mice, particularly when comparing males (Table 2 & 3), making this latter comparison the more prudent of the two.

To determine whether an age-dependent decline in retention of spatial reference memory was present for any of our three groups, the 4 time points were compared for each group separately. We then compared performance at 6–11, 12–18, and 20–24 months of age to that at 4–5 months of age, a time at which the groups did not differ. For the $APP_{SWE}$ mice, the first indication of impaired spatial reference memory was observed at 6–11 months (Fig. 3). As the mice aged, retention of spatial memory became more dramatically impaired, whereas for both the $APP_{WT}$ and Tg Neg mice, there were no age-related declines in performance (Fig. 3).

These results suggest that $APP_{SWE}$ mice do in fact have an age-dependent decline in memory but that the effect size is quite small from 6–11 months when compared to Tg Neg mice. This small effect size is most likely due to the opposing effects in $APP_{SWE}$ mice of sAPPα and AICD, which enhance cognition [16–18], and the accumulation of Aβ oligomers, which disrupt cognition. Support for this comes from comparisons between $APP_{SWE}$ and $APP_{WT}$ mice, which over-express sAPPα and AICD, but not Aβ, to the same extent as $APP_{SWE}$ mice. This comparison results in a much greater effect size, particularly at 6–11 months of age. Likewise, $APP_{WT}$ mice outperform Tg Neg mice at 6–11 and 20–24 months of age, similar to previous

reports [18]. Thus, one potential reason for disparate results in the $APP_{SWE}$ literature is the use of Tg Neg mice. When using a small number of animals with a relatively small effect size, which is often the case when experimenters compare $APP_{SWE}$ to Tg Neg mice, there is greater potential for erratic results due to variations in random sampling. By using the $APP_{WT}$ mice with their greater effect size, variability between experiments should be reduced. We should note that although we compared the Tg Negs from the $APP_{WT}$ and $APP_{SWE}$ lines to ensure that background strains were similar between the two lines, it is possible that parental strains of the two lines have diverged over a 10–15 year period and may differ from those in other laboratories. The emergence of sublines of $APP_{SWE}$ during the 15 years since the first founder was created could explain the acquisition deficits seen in mice purchased from Taconic (for example, [28–30]).

The use of $APP_{WT}$ mice, instead of Tg Neg mice, for comparison with $APP_{SWE}$ mice results in a larger difference at 6–11 months of age. Because this is the first time point at which an age-dependent decline in performance is seen (Fig. 3), the study of cognitive enhancers at this age would most likely inform therapeutic prevention in AD. Therefore, it is imperative that differences between the $APP_{SWE}$ mice and the control group be as large as possible to detect subtle enhancements in performance. The difference between the Tg Neg and $APP_{SWE}$ mice at this time-point is not statistically significant when a large cohort of animals are compared. Therefore, subtle changes in performance are likely to be undetected.

It should be noted that the $APP_{SWE}$ mouse is only a partial model of AD that lacks both neurofibrillary tangles and neuronal loss. This mouse may best be thought of as a latent AD model. If this is the case, it is not surprising that deficits are subtle and that variability is greater in this model than in those exhibiting substantial deficits at a very young age. While these characteristics make the model difficult to use, they also make it one of the best models for research on the prevention of AD. The subtle and slow progression of cognitive deficits allows for therapeutic intervention before extensive pathology is present. A key factor in the use of this model lies in understanding the small effect size when compared to Tg Neg mice and utilizing the $APP_{WT}$ mouse line to increase this effect size.

## References


1. Wenk GL. Neuropathologic changes in Alzheimer’s disease. J Clin Psychiatry 2003;64 (Suppl 9):7–10. [PubMed: 12934968]
2. Tiraboschi P, Hansen LA, Thal LJ, Corey-Bloom J. The importance of neuritic plaques and tangles to the development and evolution of AD. Neurology 2004;62:1984–9. [PubMed: 15184601]
3. Rapp MA, Reischies FM. Attention and executive control predict Alzheimer disease in late life: results from the Berlin Aging Study (BASE). Am J Geriatr Psychiatry 2005;13:134–41. [PubMed: 15703322]
4. Hsiao K. Correlative memory deficits, Aβ elevation, and amyloid plaques in trangenic mice. Science 1996;274:99–102. [PubMed: 8810256]
5. Irizarry MC, McNamara M, Fedorchak K, Hsiao K, Hyman BT. APPSwe transgenic mice develop age-related Aβ deposits and neuropil abnormalities, but no neuronal loss in CA1. J Neuropathol Exp Neurol 1997;56:965–73. [PubMed: 9291938]
6. Frautschy SA, Yang F, Irizarry M, et al. Microglial response to amyloid plaques in APPsw transgenic mice. Am J Pathol 1998;152:307–17. [PubMed: 9422548]
7. Ashe KH. Mechanisms of memory loss in Aβ and tau mouse models. Biochem Soc Trans 2005;33:591–4. [PubMed: 16042551]
8. King DL, Arendash GW, Crawford F, Sterk T, Menendez J, Mullan MJ. Progressive and gender-dependent cognitive impairment in the APP(SW) transgenic mouse model for Alzheimer’s disease. Behav Brain Res 1999;103:145–62. [PubMed: 10513583]
9. Chapman PF, White GL, Jones MW, et al. Impaired synaptic plasticity and learning in aged amyloid precursor protein transgenic mice. Nat Neurosci 1999;2:271–6. [PubMed: 10195221]

10. Pompl PN, Mullan MJ, Bjugstad K, Arendash GW. Adaptation of the circular platform spatial memory task for mice: use in detecting cognitive impairment in the APP(SW) transgenic mouse model for Alzheimer's disease. J Neurosci Methods 1999;87:87–95. [PubMed: 10065997]
11. Morgan D, Diamond DM, Gottschall PE, et al. Aβ peptide vaccination prevents memory loss in an animal model of Alzheimer's disease. Nature 2000;408:982–5. [PubMed: 11140686]
12. King DL, Arendash GW. Behavioral characterization of the Tg2576 transgenic model of Alzheimer's disease through 19 months. Physiol Behav 2002;75:627–42. [PubMed: 12020728]
13. Arendash GW, Lewis J, Leighty RE, et al. Multi-metric behavioral comparison of APPsw and P301L models for Alzheimer's disease: linkage of poorer cognitive performance to tau pathology in forebrain. Brain Res 2004;1012:29–41. [PubMed: 15158158]
14. Westerman MA, Cooper-Blacketer D, Mariash A, et al. The relationship between Aβ and memory in the Tg2576 mouse model of Alzheimer's disease. J Neurosci 2002;22:1858–67. [PubMed: 11880515]
15. Deacon RMJ, Cholerton LL, Talbot K, et al. Age-dependent and -independent behavioral deficits in Tg2576 mice. Behavioural Brain Research 2008;189:126–38. [PubMed: 18261809]
16. Ishida A, Furukawa K, Keller JN, Mattson MP. Secreted form of β-amyloid precursor protein shifts the frequency dependency for induction of LTD, and enhances LTP in hippocampal slices. Neuroreport 1997;8:2133–7. [PubMed: 9243598]
17. Meziane H, Dodart JC, Mathis C, et al. Memory-enhancing effects of secreted forms of the β-amyloid precursor protein in normal and amnestic mice. Proc Natl Acad Sci U S A 1998;95:12683–8. [PubMed: 9770546]
18. Ma H, Lesne S, Kotilinek L, et al. Involvement of β-site APP cleaving enzyme 1 (BACE1) in amyloid precursor protein-mediated enhancement of memory and activity-dependent synaptic plasticity. Proc Natl Acad Sci U S A 2007;104:8167–72. [PubMed: 17470798]
19. Laird FM, Cai H, Savonenko AV, et al. BACE1, a major determinant of selective vulnerability of the brain to amyloid-β amyloidogenesis, is essential for cognitive, emotional, and synaptic functions. J Neurosci 2005;25:11693–709. [PubMed: 16354928]
20. Lesne S, Koh MT, Kotilinek L, et al. A specific amyloid-β protein assembly in the brain impairs memory. Nature 2006;440:352–7. [PubMed: 16541076]
21. Walsh DM, Klyubin I, Fadeeva J, et al. Naturally secreted oligomers of the Alzheimer amyloid β-protein potently inhibit hippocampal long-term potential *in vivo*. Nature Biotechnology 2002;416:535–39.
22. Fitzjohn SM, Morton RA, Kuenzi F, et al. Age-related impairment of synaptic transmission but normal long-term potentiation in transgenic mice that overexpress the human APP695SWE mutant form of amyloid precursor protein. J Neurosci 2001;21:4691–8. [PubMed: 11425896]
23. Lambert M, Barlow A, Chromy B, et al. Diffusible, nonfibrillar ligands derived from Aβ1–42 are potent central nervous system neurotoxins. Proceedings of the National Academy of Science 1998;95:6448–53.
24. Hsia AY, Masliah E, McConlogue L, et al. Plaque-independent disruption of neural circuits in Alzheimer's disease mouse models. Proc Natl Acad Sci U S A 1999;96:3228–33. [PubMed: 10077666]
25. Kawarabayashi T, Younkin LH, Saido TC, Shoji M, Ashe KH, Younkin SG. Age-dependent changes in brain, CSF, and plasma amyloid (β) protein in the Tg2576 transgenic mouse model of Alzheimer's disease. J Neurosci 2001;21:372–81. [PubMed: 11160418]
26. Maei HR, Zaslavsky K, Teixeira CtM, Frankland PW. What is the most sensitive measure of water maze probe test performance? Frontiers in Integrative Neuroscience 2009:3. [PubMed: 19404411]
27. Cohen J. A power primer. Psychological Bulletin 1992;112:155–59. [PubMed: 19565683]
28. Wang J, Ho L, Zhao W, et al. Grape-Derived Polyphenolics Prevent A{β} Oligomerization and Attenuate Cognitive Deterioration in a Mouse Model of Alzheimer's Disease. J Neurosci 2008;28:6388–92. [PubMed: 18562609]
29. Ribes D, Colomina MT, Vicens P, Domingo JL. Effects of oral aluminum exposure on behavior and neurogenesis in a transgenic mouse model of Alzheimer's disease. Experimental Neurology 2008;214:293–300. [PubMed: 18834880]
30. Díaz-Ruiz C, Wang J, Ksiezak-Reding H, et al. Role of hypertension in aggravating Aβ neuropathology of AD type and tau-mediated motor impairment. Cardiovasc Psychiatry Neurol. 2009

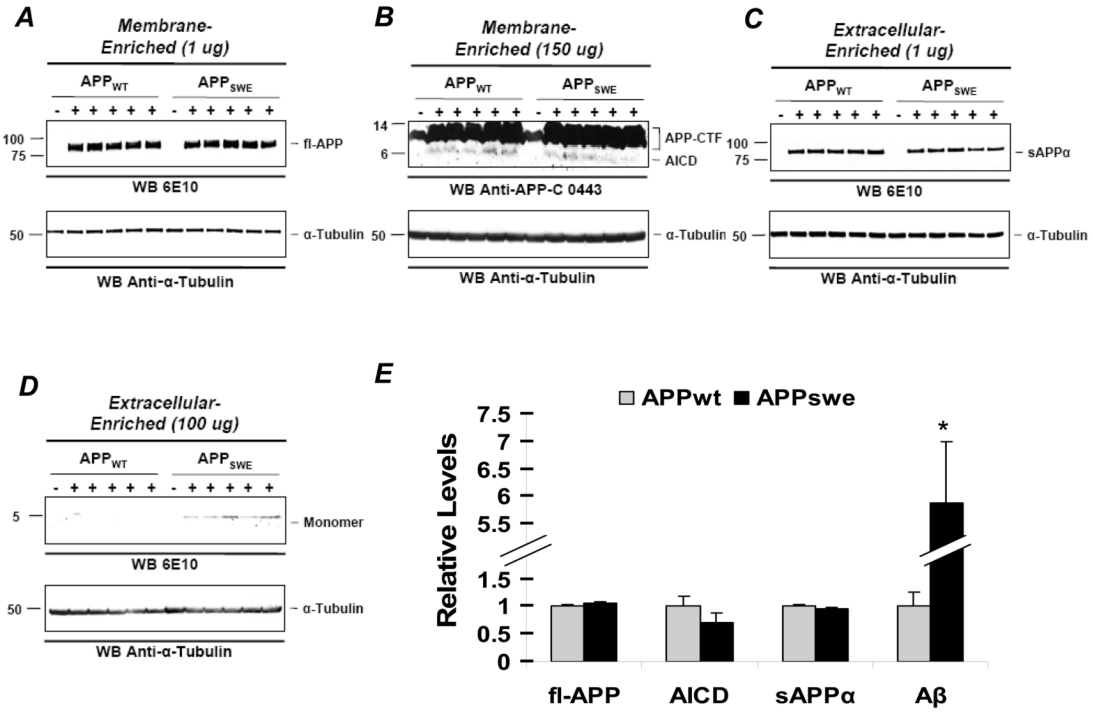


**Figure 1. Comparison of APP and APP metabolites in 8.5 month old $APP_{SWE}$ and $APP_{WT}$ mice**

***A–C,*** Forebrain lysates from 5 $APP_{WT}$ and 5 $APP_{SWE}$ mice were analyzed by SDS-PAGE and immunoblot probed with (***A***) 6E10, a mouse monoclonal antibody, to detect full-length APP (fl-APP) in the membrane-enriched fraction (MB), (***B***) a rabbit polyclonal anti-APP antibody to recognize AICD in the MB fraction, and (***C***) 6E10 to detect sAPPα in the extracellular-enriched (EC) fraction. To ensure there was no fl-APP contamination during protein extraction, the same amount of the EC fraction was immunoprecipitated with an anti-APP antibody that recognizes the C-terminal region. No fl-APP was detected from subsequent probing with 6E10 (data not shown). (***D***) 6E10 was used to detect Aβ species in the EC fraction. All blots were stripped and reprobed with anti-α-Tubulin (bottom rows). (***E***)**,** APP, AICD, sAPPα, and Aβ were quantified by normalizing the band intensity to that of α-Tubulin to determine the relative intensity between $APP_{WT}$ and $APP_{SWE}$ mice. Mean levels in $APP_{WT}$ mice were defined as 1.0. There were no significant differences in protein levels of α–Tubulin, fl-APP, AICD, or sAPPα (*P*s>0.1), but Aβ levels were significantly higher for the $APP_{SWE}$ mice ($P< 0.01$).

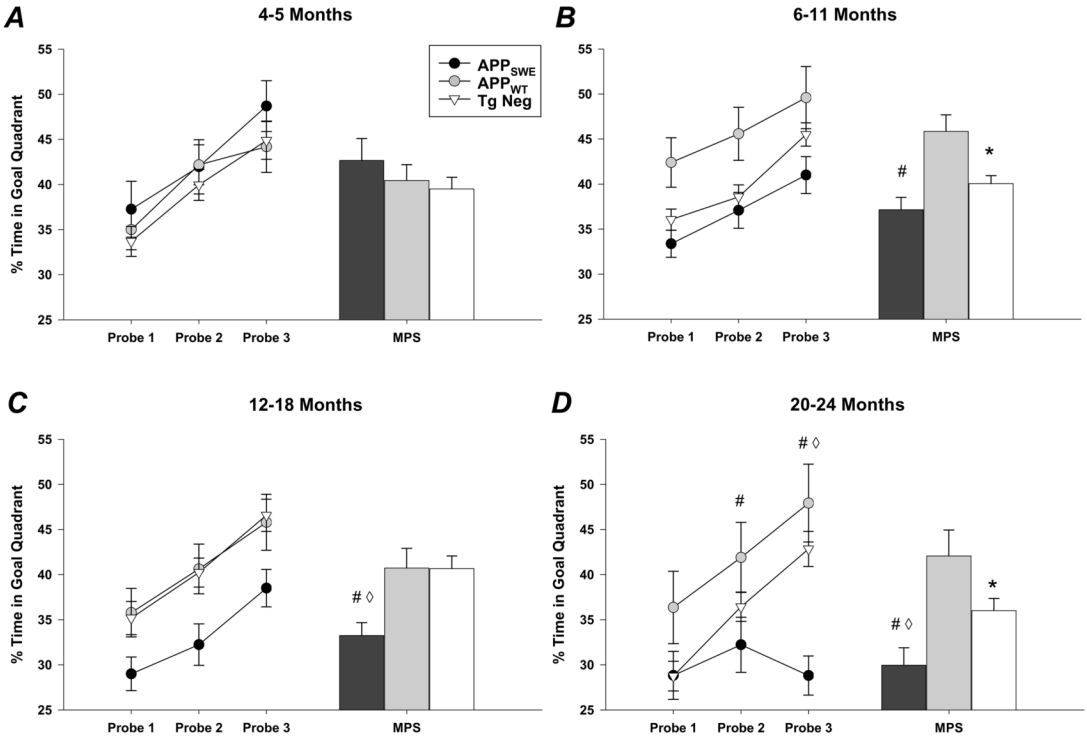


**Figure 2. Assessment of memory function in $APP_{SWE}$, $APP_{WT}$, and Tg Neg mice using the Morris water maze**

***A***, Compared with $APP_{WT}$ and Tg Negs, the time spent swimming in the target quadrant during probe trials did not differ in $APP_{SWE}$ mice at 4–5 months of age. Mean platform score (MPS) represents the average time spent in the quadrant area for the three probes conducted. RMANOVA data are as follows: transgene (MPS): $P = 0.44$; probe versus transgene: $P=0.86$. ***B***, At 6–11 months of age, $APP_{WT}$ mice spent significantly more time in the target quadrant than both $APP_{SWE}$ and Tg Neg mice. RMANOVA data are as follows: transgene (MPS): $P=0.001$; probe versus transgene: $P=0.83$. ***C***, $APP_{SWE}$ mice were impaired compared to both $APP_{WT}$ and Tg Neg mice. RMANOVA data are as follows: transgene (MPS): $P=0.003$; probe versus transgene: $P=0.97$. ***D***, Although $APP_{WT}$ and Tg Neg mice improved with repeated probe testing, $APP_{SWE}$ mice exhibited stable performance (probe trial RMANOVA: $P=0.46$), resulting in a probe by transgene interaction. RMANOVA data are as follows: transgene (MPS): $P=0.001$; probe versus transgene: $P=0.004$. # APPSWE vs. APPWT $p<0.05$, APPSWE vs. Tg Neg $p<0.05$, * APPWT vs. Tg Neg $p<0.05$.

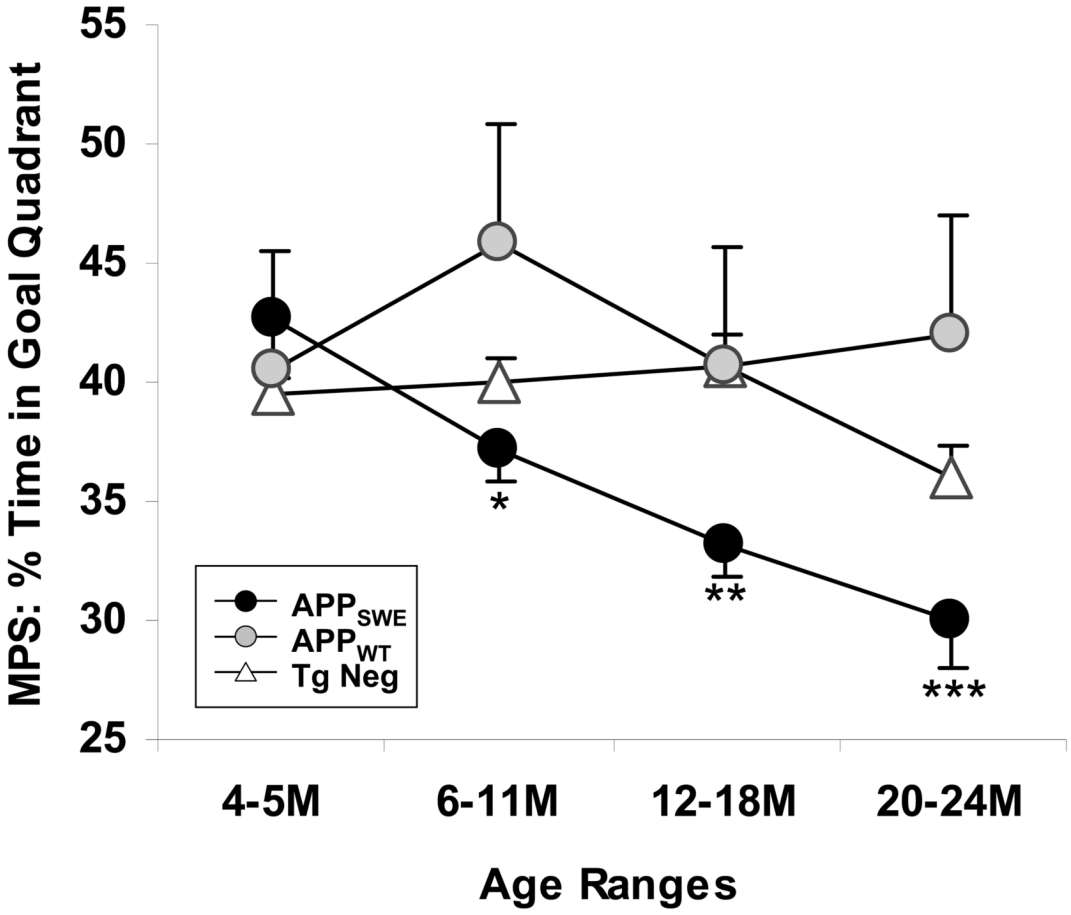


**Figure 3. $APP_{SWE}$ mice develop age-dependent memory deficits not seen in $APP_{WT}$ or Tg Neg mice**

Mean platform score (MPS) represents the average time spent in the quadrant area for the three probes conducted.. $APP_{SWE}$ mice exhibited an age-dependent decline in MPS performance ($P$=0.0002) that was not observed in $APP_{WT}$ ($P$=0.23) or Tg Neg mice ($P$=0.07). 4–5M vs. older ages *$p < 0.05$, **$p < 0.001$, ***$p < 0.0001$.

## Table 1

Animals tested in the MWM.

| | Included | |
|---|---|---|
| **Mice Tested** | **Male** | **Female** |
| $APP_{WT}$ (Tg5469) | | |
| 4–5 Months | 11 | 13 |
| 6–11 Months | 14 | 8 |
| 12–18 Months | 9 | 19 |
| 20–24 Months | 6 | 6 |
| $APP_{SWE}$ (Tg2576) | | |
| 4–5 Months | 10 | 15 |
| 6–11Months | 29 | 33 |
| 12–18 Months | 20 | 16 |
| 20–24 Months | 15 | 8 |
| Tg neg | | |
| 4–5 Months | 23 | 25 |
| 6–11Months | 41 | 44 |
| 12–18 Months | 30 | 37 |
| 20–24 Months | 23 | 20 |
| Total | 231 | 244 |

**Table 2**

Effect and sample sizes of $APP_{SWE}$ and $APP_{WT}$ B6/SJL mice tested in the Morris water maze.

| | | Males and Females Combined $APP_{SWE}$ vs. Tg Neg | | | | Males and Females Combined $APP_{SWE}$ vs. $APP_{WT}$ | | | | Males and Females Combined $APP_{WT}$ vs. Tg Neg | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Age** | **Probe** | **Mean Difference[1]** | **Pooled Standard Deviation** | **Effect Size** | **N Needed per Group** | **Mean Difference[1]** | **Pooled Standard Deviation** | **Effect Size** | **N Needed per Group** | **Mean Difference[2]** | **Pooled Standard Deviation** | **Effect Size** | **N Needed per Group** |
| 4–5 Months | Probe 1 | −3.57 | 12.79 | 0.28 | 203 | −2.27 | 13.16 | 0.17 | 528 | −1.30 | 11.11 | 0.12 | 1146 |
| | Probe 2 | −2.00 | 12.87 | 0.16 | 650 | 0.22 | 12.89 | 0.02 | 53885 | −2.22 | 11.41 | 0.20 | 416 |
| | Probe 3 | −3.81 | 14.11 | 0.27 | 216 | −4.51 | 13.70 | 0.33 | 146 | 0.70 | 14.06 | 0.05 | 6330 |
| | Mean Probe | −3.13 | 10.01 | 0.31 | 162 | −2.19 | 10.33 | 0.21 | 350 | −0.94 | 8.71 | 0.11 | 1347 |
| 6–11 Months | Probe 1 | 2.68 | 11.16 | 0.24 | 273 | 9.03 | 11.95 | 0.76 | 29 | −6.35 | 11.14 | 0.57 | 50 |
| | Probe 2 | 1.49 | 13.70 | 0.11 | 1326 | 8.49 | 15.02 | 0.57 | 51 | −7.00 | 12.49 | 0.56 | 51 |
| | Probe 3 | 4.51 | 13.77 | 0.33 | 148 | 8.60 | 15.96 | 0.54 | 55 | −4.09 | 12.78 | 0.32 | 155 |
| | Mean Probe | 2.90 | 9.32 | 0.31 | 163 | 8.71 | 10.18 | 0.86 | 23 | −5.81 | 8.19 | 0.71 | 33 |
| 12–18 Months | Probe 1 | 6.18 | 13.70 | 0.45 | 79 | 6.78 | 12.36 | 0.55 | 54 | −0.60 | 14.66 | 0.04 | 9368 |
| | Probe 2 | 7.99 | 13.26 | 0.60 | 45 | 8.38 | 13.93 | 0.60 | 45 | −0.39 | 13.46 | 0.03 | 18695 |
| | Probe 3 | 5.70 | 13.86 | 0.41 | 94 | 7.31 | 14.07 | 0.52 | 60 | −1.61 | 15.10 | 0.11 | 1380 |
| | Mean Probe | 5.40 | 10.28 | 0.53 | 58 | 7.49 | 9.84 | 0.76 | 29 | −2.09 | 11.18 | 0.19 | 450 |
| 20–24 Months | Probe 1 | −0.08 | 11.32 | 0.01 | 314300 | 7.53 | 12.79 | 0.59 | 47 | −7.61 | 11.27 | 0.68 | 36 |
| | Probe 2 | 4.21 | 11.98 | 0.35 | 128 | 9.68 | 13.84 | 0.70 | 34 | −5.47 | 11.09 | 0.49 | 66 |
| | Probe 3 | 14.03 | 11.84 | 1.19 | 13 | 19.11 | 11.74 | 1.63 | 8 | −5.08 | 13.03 | 0.39 | 105 |
| | Mean Probe | 6.05 | 8.85 | 0.68 | 35 | 12.11 | 9.24 | 1.31 | 11 | −6.06 | 8.91 | 0.68 | 35 |

[1] A negative value indicates a higher mean for APP SWE (Tg2576)

[2] A negative value indicates a higher mean for APP WT (Tg5469)

**Table 3**

Gender effects on effect and sample size of $APP_{SWE}$ and $APP_{WT}$ B6/SJL mice tested in the Morris water maze.

| | | Males Only $APP_{SWE}$ vs. Tg Neg | | | | Males Only $APP_{SWE}$ vs. $APP_{WT}$ | | | | Males Only $APP_{WT}$ vs. Tg Neg | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Age** | **Probe** | **Mean Difference[1]** | **Pooled Standard Deviation** | **Effect Size** | **N Needed per Group** | **Mean Difference[1]** | **Pooled Standard Deviation** | **Effect Size** | **N Needed per Group** | **Mean Difference[2]** | **Pooled Standard Deviation** | **Effect Size** | **N Needed per Group** |
| 4–5 Months | Probe 1 | −3.89 | 10.48 | 0.37 | 115 | −4.84 | 12.74 | 0.38 | 110 | 0.95 | 9.99 | 0.10 | 1735 |
| | Probe 2 | −6.15 | 11.94 | 0.52 | 61 | 0.97 | 9.44 | 0.10 | 1486 | −7.12 | 11.61 | 0.61 | 43 |
| | Probe 3 | −6.61 | 12.35 | 0.54 | 56 | −10.32 | 12.41 | 0.83 | 24 | 3.70 | 13.41 | 0.28 | 207 |
| | Mean Probe | −5.55 | 8.71 | 0.64 | 40 | −4.73 | 8.33 | 0.57 | 50 | −0.82 | 9.17 | 0.09 | 1961 |
| 6–11 Months | Probe 1 | 2.37 | 11.66 | 0.20 | 381 | 11.79 | 13.51 | 0.87 | 22 | −9.42 | 10.89 | 0.87 | 22 |
| | Probe 2 | 0.25 | 13.87 | 0.02 | 48315 | 6.58 | 15.32 | 0.43 | 86 | −6.33 | 12.25 | 0.52 | 60 |
| | Probe 3 | 2.20 | 13.79 | 0.16 | 616 | 5.63 | 17.44 | 0.32 | 152 | −3.44 | 12.81 | 0.27 | 219 |
| | Mean Probe | 1.61 | 9.74 | 0.17 | 574 | 8.00 | 11.53 | 0.69 | 34 | −6.39 | 7.69 | 0.83 | 24 |
| 12–18 Months | Probe 1 | 11.43 | 12.15 | 0.94 | 19 | 13.30 | 13.51 | 0.98 | 18 | −1.86 | 14.85 | 0.13 | 1000 |
| | Probe 2 | 11.33 | 13.04 | 0.87 | 22 | 12.12 | 15.82 | 0.77 | 28 | −0.79 | 13.12 | 0.06 | 4326 |
| | Probe 3 | 8.09 | 13.30 | 0.61 | 44 | 6.36 | 14.23 | 0.45 | 80 | 1.73 | 15.03 | 0.12 | 1184 |
| | Mean Probe | 10.28 | 9.81 | 1.05 | 16 | 10.59 | 10.67 | 0.99 | 17 | −0.31 | 12.36 | 0.03 | 24951 |
| 20–24 Months | Probe 1 | −1.33 | 10.29 | 0.13 | 939 | 18.71 | 11.65 | 1.61 | 8 | −20.04 | 9.30 | 2.16 | 5 |
| | Probe 2 | 9.07 | 10.79 | 0.84 | 24 | 21.23 | 13.18 | 1.54 | 8 | −12.16 | 9.79 | 1.24 | 12 |
| | Probe 3 | 14.41 | 11.75 | 1.23 | 12 | 18.04 | 10.95 | 1.65 | 7 | −3.63 | 12.00 | 0.30 | 173 |
| | Mean Probe | 7.38 | 8.04 | 0.92 | 20 | 19.33 | 8.11 | 2.39 | 4 | −11.94 | 8.19 | 1.46 | 9 |

| | | Females Only $APP_{SWE}$ vs. Tg Neg | | | | Females Only $APP_{SWE}$ vs. $APP_{WT}$ | | | | Females Only $APP_{WT}$ vs. Tg Neg | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Age** | **Probe** | **Mean Difference[1]** | **Pooled Standard Deviation** | **Effect Size** | **N Needed per Group** | **Mean Difference[1]** | **Pooled Standard Deviation** | **Effect Size** | **N Needed per Group** | **Mean Difference[2]** | **Pooled Standard Deviation** | **Effect Size** | **N Needed per Group** |
| 4–5 Months | Probe 1 | −3.54 | 14.38 | 0.25 | 260 | −0.29 | 13.38 | 0.02 | 33412 | −3.25 | 11.95 | 0.27 | 213 |
| | Probe 2 | 1.41 | 13.27 | 0.11 | 1389 | −0.70 | 14.79 | 0.05 | 7004 | 2.11 | 10.75 | 0.20 | 408 |
| | Probe 3 | −2.22 | 15.16 | 0.15 | 732 | −0.30 | 14.19 | 0.02 | 35117 | −1.92 | 14.50 | 0.13 | 895 |

| | | Females Only $APP_{SWE}$ vs. Tg Neg | | | | Females Only $APP_{SWE}$ vs. $APP_{WT}$ | | | | Females Only $APP_{WT}$ vs. Tg Neg | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Age** | **Probe** | **Mean Difference[1]** | **Pooled Standard Deviation** | **Effect Size** | **N Needed per Group** | **Mean Difference[1]** | **Pooled Standard Deviation** | **Effect Size** | **N Needed per Group** | **Mean Difference[2]** | **Pooled Standard Deviation** | **Effect Size** | **N Needed per Group** |
| | Mean Probe | −1.45 | 10.86 | 0.13 | 880 | −0.43 | 11.48 | 0.04 | 11185 | −1.02 | 8.28 | 0.12 | 1034 |
| 6–11 Months | Probe 1 | 2.96 | 10.69 | 0.28 | 206 | 3.98 | 9.62 | 0.41 | 93 | −1.03 | 11.11 | 0.09 | 1825 |
| | Probe 2 | 2.64 | 13.50 | 0.20 | 411 | 11.71 | 14.59 | 0.80 | 26 | −9.07 | 12.57 | 0.72 | 32 |
| | Probe 3 | 6.53 | 13.57 | 0.48 | 69 | 11.47 | 13.90 | 0.83 | 25 | −4.94 | 12.73 | 0.39 | 106 |
| | Mean Probe | 4.04 | 8.88 | 0.46 | 77 | 9.05 | 8.45 | 1.07 | 15 | −5.01 | 8.69 | 0.58 | 49 |
| 12–18 Months | Probe 1 | 2.58 | 13.32 | 0.19 | 419 | 4.85 | 10.06 | 0.48 | 69 | −2.27 | 12.32 | 0.18 | 463 |
| | Probe 2 | 4.99 | 13.18 | 0.38 | 111 | 6.10 | 11.97 | 0.51 | 62 | −1.11 | 13.35 | 0.08 | 2268 |
| | Probe 3 | 9.35 | 13.62 | 0.69 | 35 | 9.99 | 13.50 | 0.74 | 30 | −0.64 | 14.68 | 0.04 | 8256 |
| | Mean Probe | −4.36 | 10.17 | 0.43 | 87 | 6.98 | 8.58 | 0.81 | 25 | −11.34 | 9.58 | 1.18 | 13 |
| 20–24 Months | Probe 1 | 0.15 | 11.80 | 0.01 | 97141 | −5.10 | 10.92 | 0.47 | 73 | 5.26 | 10.69 | 0.49 | 66 |
| | Probe 2 | −3.56 | 12.66 | 0.28 | 200 | −4.50 | 11.32 | 0.40 | 101 | 0.94 | 10.64 | 0.09 | 2009 |
| | Probe 3 | 13.79 | 11.93 | 1.16 | 13 | 20.43 | 12.80 | 1.60 | 8 | −6.64 | 14.05 | 0.47 | 72 |
| | Mean Probe | 3.46 | 9.68 | 0.36 | 124 | 3.61 | 8.93 | 0.40 | 97 | −0.15 | 8.83 | 0.02 | 54394 |

[1] A negative value indicates a higher mean for $APP_{SWE}$ (Tg2576)

[2] A negative value indicates a higher mean for $APP_{WT}$ (Tg5469)