A DAY IN THE LIFE OF MUS MUSCULUS: HOMECAGE BEHAVIOURAL ANALYSIS OF A MOUSE MODEL OF ALZHEIMER DISEASE

JIAJIE LIANG
Bachelor of Science, University of Lethbridge, 2020

A thesis submitted
in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in

NEUROSCIENCE

Department of Neuroscience
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

A DAY IN THE LIFE OF MUS MUSCULUS: HOMECAGE BEHAVIOURAL ANALYSIS OF A MOUSE MODEL OF ALZHEIMER DISEASE

JIAJIE LIANG

Date of Defence: December 11, 2023

Dr. M. Mohajerani	Associate Professor	Ph.D.
Dr. H. Cheng	Associate Professor	Ph.D.
Thesis Co-Supervisors		
Dr. I. Whishaw	Professor	Ph.D.
Thesis Examination Committee Member		
Dr. A. Luczak	Professor	Ph.D.
Thesis Examination Committee Member		
Dr. D. Euston	Associate Professor	Ph.D.
Chair, Thesis Examination Committee		

Dedication

To all who have supported, motivated, and inspired me along the way.

Abstract

Traditionally, behavioural research of Alzheimer disease animal models involved specialized experiments that require dedicated apparatus and presence of experimenter. However, experiment apparatus and the interaction between animals and experimenters can influence the behaviours of the animals, and result in difficulty in reproducibility. One recent innovation is to study behaviours of Alzheimer disease mouse model in their homecages. This thesis presents an experiment using automated homecages to observe the homecage behaviours of 5xFAD mouse models over 26 weeks. By measuring daily activity level, circadian rhythm and excursion behaviours, the experiment successfully produces measurements consistent with prior knowledge and provides some further insight in the behaviours of the mouse model. This thesis validates the approach using homecage behaviours as a paradigm for AD animal research.

Acknowledgement

I would like to take this opportunity to express my deep appreciation and gratitude to the following individuals and organizations whose support and assistance have been invaluable in the completion of this thesis:

My supervisors, Dr. M. Mohajerani and Dr. H. Cheng: I am grateful to my supervisors, Dr. M. Mohajerani and Dr. H. Cheng for their support, guidance and mentorship throughout this research journey.

My supervisory committee: I extend my sincere thanks to other members of my supervisory committee, Dr. I. Whishaw and Dr. A. Luczak for their valuable insights and constructive feedback.

Homecage hardware team led by Dr. H. Ryait: I would like to acknowledge all current and previous members in the homecage hardware team in Mohajerani lab for providing the support for the automated homecage system.

Homecage software team led by Dr. E. Bermudez-Contreras and B. Penner: I would like to acknowledge all current and previous members in the homecage software team in Mohajerani lab for the collaboration in building and supporting the automated homecage software system.

Colleagues: I am thankful to all my colleagues providing me support and advice during my research journey. I am especially thankful to Dr. E. Bermudez-Contreras for his support and mentorship.

Alberta Innovate: This research was made possible through the financial support provided by Alberta Innovate.

Animal Care Services: I would like to acknowledge the Animal Care Services staff of University of Lethbridge for providing access to the state-of-the-art faccilities and animal care supports in the duration of the experiment.

Table of Contents

List of Figures

List of Abbreviations

AD	Alzheimer disease
APO	apolipoprotein E
APP	amyloid precursor protein
Aβ	beta-amyloid
BRAC	basic rest-activity cycle
CCAC	Canadian Council on Animal Care
CR	conditional response
CS	conditional stimulus
DLC	DeepLabCut
EMG	electromyography
EOAD	early-onset Alzheimer disease
fAD	sporadic Alzheimer disease
FPS	frames per second
GPU	Graphics Processing Unit
IoU	intersection over union
LFP	local field potential
LOAD	late-onset Alzheimer disease
mAP	mean average precision
MoSeq	Motion Sequencing
MWM	Morris water maze
NFT	neurofibrillary tangle
NHP	Non-human primate
NPS	neuropsychiatric symptom
NREM	Non-rapid eye movement
PR	precision-recall
PSEN	presenilin
RFID	Radio Frequency Identification
RPi	RaspberryPi
sAD	familial Alzheimer disease
SORT	Simple Online and Realtime Tracking
TB	terabytes
UR	unconditional response
US	unconditional stimulus
VWR	voluntary wheel running
YOLO	You Only Look Once

Introduction

Can insight into a cure for Alzheimer disease (AD) be obtained from the study of mice? This is the question addressed in the present thesis. The following sections will first describe what AD is, followed by the mouse models of AD, and finally common behavioral assessments of normal and Alzheimer mouse models. This background will provide insights for the design of the experiments and results reported in the thesis.

Alzheimer Disease

Dementia is the progressive impairment of cognitive function, and it is estimated that over 55 millions of people are living with dementia around the world currently (World Health Organization, 2021). It is one of the leading causes of dependency, disability, and death, and it is by far the fastest growing cause since year 2000. AD, a progressive neurodegenerative disorder, is the leading cause of dementia and accounts for over 50% of the cases (Lane, Hardy, & Schott, 2018). While different risk factors for AD have been identified (Xu, et al., 2015), there are currently no established cure for the disease.

Based on the onset, AD can be divided into early-onset AD (EOAD), which has onset before the age of 65 and accounts for 1% to 6% of all cases, and late-onset AD (LOAD), which has onset after age of 65 (Piaceri, Nacmias, & Sorbi, 2013). Genetically, AD can be divided into familial AD (fAD), which is predominately early-onset, and sporadic AD (sAD), which shows no family link. The vast majority of cases of AD are late-onset and sporadic. Familial AD usually results from mutations in one of the three

genes - amyloid precursor protein (APP), presenilin 1 (PSEN1) and presenilin 2 (PSEN2). On the other hand, the sporadic form of the disease is from more complex interaction between genes and environmental factors, with the apolipoprotein E (APOE) gene being considered the biggest risk factor (Lane, Hardy, & Schott, 2018).

Two common features of AD are amyloid plaques and neurofibrillary tangles (NFTs), and these processes cause synaptic and neuronal loss (Lane, Hardy, & Schott, 2018). Amyloid plaques are extracellular deposits of abnormally folded beta-amyloid (Aβ) with 40 or 42 amino acids (Aβ40 and Aβ42), which are byproducts of APP cleavage by β- and γ-secretase enzymes in the brain. Meanwhile, NFTs are intraneuronal accumulation of paired helical filaments consisting of abnormally phosphorylated tau proteins in the brain (Bancher, et al., 1989). Amyloid plaques appear and plateau early as the disease progresses to the symptomatic phase, while tangle formation parallels neuronal and synapse loss (Lane, Hardy, & Schott, 2018).

The leading theory of AD pathogenesis is the amyloid hypothesis (Lane, Hardy, & Schott, 2018). The hypothesis suggests that AD primarily stems from the accumulation of abnormal forms of Aβ due to an imbalance between Aβ production and Aβ clearance. Formation of NFTs, loss of neuron and synapses are downstream processes within this hypothesis. This is supported by the fact that many genetic mutations recognized to be major risk factors of AD are linked to the amyloid processes. Specifically, all three fAD mutations are involved in Aβ generation or processing, while APOE in sAD plays a role in amyloid clearance. However, it is noteworthy that there are elderly individuals passing with significant β-amyloid depositions but no symptoms of AD. This corroborates theory

suggesting that soluble Aβ oligomers may be the more toxic forms, while the plaque may be a protective mechanism against toxic Aβ. These oligomers can also lead to hyperphosphorylation of tau, which lead to NFTs. While tau pathology is closely associated to the process of AD, tau mutations alone do not lead to AD. The mechanical connection between amyloid and tau pathology in pathogenesis of AD is still unclear despite their clear significance.

Animal Models of Alzheimer Disease

To better study and understand AD, animals are used in experiments not suitable for human participants. Due to AD not being a part of normal aging, animal models that can acquire AD need to be developed. Non-human primates (NHPs) are frequently employed in neuroscience research due to the similarities between their brains and human brains, as well as their ability to perform various perceptual and cognitive tasks akin to those used in human studies (De Felice & Munoz, 2016). However, due to their long lifespans, development of AD model can take decades. Laboratory mice (Mus musculus) and laboratory rats (Rattus norvegicus domestica) are much more commonly used in AD research, and each has their advantages. Lab rats, due to their bigger size and brain, can be trained and used in more complex behaviours and provide larger tissue samples. Lab mice, on the other hand, are genetically closer to human. This makes them more suitable for using in both genetic manipulation studies and creation of genetically manipulated animal models for other research. Lab mice were used in the experiments in this thesis.

There exists many genetically modified mouse models of AD, depending on the genetic alterations they receive. Several common approaches are transgenic, knock-in and

knock-out. Transgenic mouse models are created by introducing exogenous genes, typically human genes associated with AD, into mouse genome. Examples of transgenic mouse models include 5xFAD transgenic mice, which carry 5 fAD mutations in the APP and PSEN1 genes. 5xFAD mice exhibit rapid and aggressive amyloid plaque formation and start displaying symptoms at as early as 6 months old. Tau transgenic mouse models overexpress mutated tau protein, leading to formation of NFTs without Aβ pathology. 3xTg-AD mice contain mutations in APP, PSN1 and Tau genes and exhibit both amyloid plaques and NFTs.

In knock-in models, genetic alterations, such as mutations in genes, are introduced into the mouse genome. As a result, no exogenous genes are introduced into the genome. Some examples of knock-in mouse models are various APP knock-in mouse models, which incorporate fAD mutations into mouse APP genes. For example, APP-NL-F mouse model includes two mutations in APP and starts showing cognitive impairments at 18 months, while APP-NL-GF mouse model includes an additional mutation which lead to much earlier cognitive impairments at age of 6 months (Kundu, et al., 2021).

Due to the different approaches involved for transgenic mouse models and knock-in mouse models receiving the mutations of interest, extra considerations are needed when selecting the mouse models. Knock-in mouse models preserve the endogenous expression patterns, and it may be intuitively more preferrable (Jankowsky & Zheng, 2017). However, this may not always be possible. Meanwhile transgenic mouse models may have the tradeoff of unexpected confounding effects from introduction of exogenous

genes, there are still situations where transgenic models do better, especially in expressing disease pathology.

Instead of receiving exogenous genes or mutations into mouse gene, knock-out mouse models have specific genes disrupted, leading to loss or inactivity of the corresponding protein products. These mouse models are used to investigate the functions of specific gene by eliminating it. For example, APP knock-out mice are used to explore the role of APP gene in Aβ production and APOE knock-out mice are used to investigate the role of APOE gene in Aβ clearance. Beyond the three approaches, there are mouse models created through gene delivery with viral vector and chemicals or drugs inductions (Van Dam & Deyn, 2011).

Assessments of Behaviours of AD Mouse Models

As progressive loss of neurons and synapses from Alzheimer disease leads to continuous deterioration of cognitive ability, behavioural assessments are effective ways of evaluating the progress of the disease without involving etiopathogenesis of the disease (Puzzo, Lee, Palmeri, Calabrese, & Arancio, 2014). Many methods of behaviour assessments have been applied to research with AD mouse models. This section will describe some common approaches.

Fear conditioning is a behavioural task assessing the associative memory, which is affected in AD patient. Conditioning tasks test associative learning between two stimuli – unconditional stimulus (US), which can evoke a typical behavioural response, and conditional stimulus (CS), which is a neutral stimulus with no relation to the response

being learned. The training process involves repeatedly presenting US and CS together or in quick succession as US trigger the subject's response (unconditional response – UR) (Puzzo, Lee, Palmeri, Calabrese, & Arancio, 2014). The learning is successful when CS alone can trigger the subject's UR associated with US (conditional response – CR). In one common example of fear conditioning task for mice, the US is electric shocks, and CS can be audible ringtones. UR in this case is the automatic reactions triggered by the unpleasant experience of electric shocks, and one of the reactions can be freezing in motion. By presenting the ringtones and the electric shocks together, the mice will eventually show reactions such as freezing when only the ringtone is delivered (CR). The severity of AD then can be assessed through the robustness of the CR. Multiple regions of the brain are involved in the fear conditioning tasks. Amygdala is engaged in the fear response and formation and storage of emotional memories. Hippocampus is responsible for learning the context that triggers fear and consolidating the memory, while neocortex stores the memory in the end. As sensory inputs play a role in fear conditioning task, thalamus is also involved.

Morris water maze (MWM) is a behavioural task testing hippocampal-dependent spatial learning and long-term spatial memory (Morris, 1981). While originally designed for rats, it has also been applied to mice models of AD (Puzzo, Lee, Palmeri, Calabrese, & Arancio, 2014). MWM uses a large circular pool filled with opaque water, a submerged small platform just beneath the surface of water and potentially some objects as cues around the pool. MWM consists of three tests – training, probe trials and visible tests. The training last for 2 to 10 days. During training, the mouse individually goes through one or multiple daily training sessions where they are placed in the pool and must search for the

hidden platform. Their trials of learning and locating the platforms are recorded. The time the mice need and the length of path they take to reach the platform should both progressively decrease as they learned the location of the platform. Probe trials happen 24 hours after training. In probe trials, the submerged platform is removed. The mice are individually released into the water to freely swim around the pool. The entire swimming session is recorded, and the time the mice are swimming around the quadrant, where the platform is originally, is recorded. Longer time swimming and seeking for the platform around said quadrant shows more robust spatial memory. In the end, visible tests are done to ensure the mice haven't developed motivational, visual or motor impairments during the experiment period. The mice are placed in the pool where the hidden platform is marked with visible cue, and the performance of them reaching the platform is recorded. Normal performance on visible test but decreased performance in training is an indication of impairment spatial memory. As MWM mainly tests spatial memory, it relies on the hippocampus and other brain regions that affect spatial navigation and motor skill.

Another memory related test is object recognition test, which operates on the principle that animals prefer exploring novel objects. As a result, between novel objects and familiar objects, mice would spend more time exploring the novel objects. In the training phase, mice individually are placed in a testing box containing two objects (Mehla, et al., 2019). The mice familiarize themselves with the objects (familiar objects), and then are returned to their home cage. The time the mice exploring the objects is recorded. Twenty-four hours after training, the mice are placed in a testing box with one familiar object and a novel object. The time the mice spend exploring between the familiar object and the novel object is recorded. By comparing the time spent between the

objects, the integrity of the mice object-associated memory can be assessed. Object recognition memory depends on the perirhinal cortex circuitry, which is assessed in this behavioural task (Antunes & Biala, 2012).

Open field test can also be used to assess novelty-induced behaviours, but it can also measure locomotive function, anxiety-like and exploratory behaviours (Mehla, et al., 2019). Locomotor ability is important for assessing well-being and physical health of animal in general, and increased anxiety-like behaviours can be a sign of disrupted neural circuits. During the test the mice are individually place in a large open-field box, and their locomotor activities are recorded and analyzed. Metrics such as moving distance, time in the centre of the field, number of rearing and numbers of fecal boli are often recorded and analyzed (Sturman, Germain, & Bohacek, 2018). More moving distance, more rearing, and more time spent in the centre are all sign of increased exploration and lower stress.

A common symptom for AD in human is disrupted sleep. As a result, circadian rhythm is another common research area for AD mouse model. While there are invasive methods to measure the sleep and wake states of the animal such as local field potential (LFP) and electromyography (EMG) recording from the hippocampus (Singh, Bermudez-Contreras, Nazari, Sutherland, & Mohajerani, 2019), there are also behavioural approaches for it. One common way to measure the circadian rhythm of mice is through voluntary wheel running (VWR). This is through providing running-wheels in the home cages of the animals and recording the usage of the running-wheels. As animals voluntarily using the running wheel is a robust phenomenon, this provides a good approximation of the animal activity schedule (Sherwin, 1998). Other methods of

measuring mouse activity in home cage involve using specialize hardware such as infra-red motion detector at certain position of the cage, or piezoelectric system under the bottom of the cage that detects pressure changes (Duncan, et al., 2012). Analyses of the estimated activity usually focus on periodogram power, average activity in the light and dark cycle, the activity onset at the transitions of the cycles, and the activity bouts (Brown, Fisk, Pothecary, & Peirson, 2019).

Problems with Assessments

Last section described several behavioural assessment methods for animal models of AD, but there are many more. These behavioural tasks are important as they target very specific brain functions and/or behavioural characteristics. However, there can be some problems in the usage of these methods. This section will explain some of these potential issues.

One potential problem of many behavioural tasks is that they are inherently unnatural to the mice. Mice are nocturnal animals, so they are more active in their night cycle while spending the majority of light cycle resting. However, most of the behavioural tasks happens in room with bright light during the light cycle of the mice, and illumination can affect the behaviours of the mice (Farnworth, Innes, & Waas, 2016). In addition, the presence of the experimenters and being handled by experimenters can create stress on the animals (Faraji, et al., 2022). Furthermore, mice are territorial animals. The experiment setups usually do not have odours of the mice, this can have behavioural effects on the mice (Gray & Hurst, 1995). There is also the issue where some tasks are designed for lab rats then applied to lab mice. For example, lab mice perform

worse in MWM than lab rats do as lab mice are not as good at swimming. Also, for VWR specifically, the usage of running wheel has an effect to the mice health and behaviours due to the extra exercise from wheel running (Ramanathan, Stowie, Smale, & Nunez, 2010). While these factors can be controlled in the experiments, they can still be confounding factors in the behaviours being observed.

Another potential problem is reproducibility. While these behavioural tasks all have standardized protocols, it is not a guarantee of reproducibility. Studies have shown that experiments of same behavioural paradigms carrying out in different labs can yield different results (Crabbe, Wahlsten, & Dudek, 1999). Experimenter is one factor that is difficult to control. Some behavioural tasks rely on human scoring. Human scoring is not only labour intensive, but also can lead to bias as behaviours are defined and described intuitively (Berman, 2018). In addition, different experimenters' handling can induce different level of stress (Chesler, Wilson, Lariviere, Rodriguez-Zas, & Mogil, 2002). Experiments have shown that factors such as sex of the experimenters, and familiarity of the animals to the experiments can have effects on stress level and behaviours of the animals (Faraji, et al., 2022).

Thesis Proposal

To address the problems described above, this thesis proposes using automated homecage monitoring of behaviours (Bermudez-Contreras, Sutherland, Mohajerani, & Whishaw, 2022) as a behavioural paradigm for mouse models of AD. Using standardized automated homecages as housing for mouse models allows long-term continuous recording and monitoring of their behaviours with minimal human interference (Hobson,

Bains, Greenaway, Wells, & Nolan, 2020; Iannello, 2019; De Chaumont, et al., 2019). This paradigm allows recording of the most natural behaviours of the mice while minimizing environmental variation to minimize the reproducibility issues (Richter, 2020; Richardson, 2015). The concept of reaction norm, introduced by Richard Woltereck over 100 years ago, suggests that phenotypic characteristics can vary in response to environmental variations for a single genotype (Voelkl & Würbel, 2016). Similarly, different genotypes within a species can display different phenotypic expressions in the same environment. This thesis hypothesizes that homecage behaviours, as phenotypic traits, can be used to detect behavioural differences between 5xFAD AD mouse models and its non-transgenic counterparts in the same environment – the automated homecages. This thesis aims to use long-term homecage recording to detect the divergence of activity level and circadian rhythm as the mice age. In addition, the thesis aims to use excursion behaviours (Eilam & Golani, 1989) within the homecage to assess exploratory behaviours.

Author's Contribution

The author of this thesis had made the following contributions to the thesis. The author spent in total 2 and a half years working in the software team to design, implement, and test the homecage monitoring system, partly concurrent with the experiments. An 8-month long pilot experiment, similar to the experiment showcased in the thesis, was conducted by the author to both test the systems and prepare for the main experiment. The author conducted the experiment showcased in the thesis and analyzed the data.

Materials and Method

This section will introduce the animals, apparatuses and procedures involved in the experiment, and the analyses done to the data generated.

Animals

The experiment used 6 female 5xFAD (B6SJL) mice as the experiment group, and 6 non-transgenic female littermates as the control group. All mice were moved to the automated homecages in a dedicated room and singly housed at the age of 15 to 16 weeks. The room was set to a 12h/12h light and dark cycle to mimic a stable day-night cycle, and lights came on at 7:30am and turned off at 7:30pm. Mice had ad lib access to food and water.

Apparatus

The experiment used 24 automated homecages from Homecage Monitoring System (Singh, Bermudez-Contreras, Nazari, Sutherland, & Mohajerani, 2019). All cages used were the standard Optimice blockparty cages with customized tops for the monitoring hardware. Cages were spread out over 2 double-layer circular racks. Half of the cages served as the homecages and equipped with food hopper and water bottle, and sawdust as bedding material. Crinkle paper and nestlets were provided for nest building. The other 12 cages (side cages) were set between the homecages and only contained sawdust.

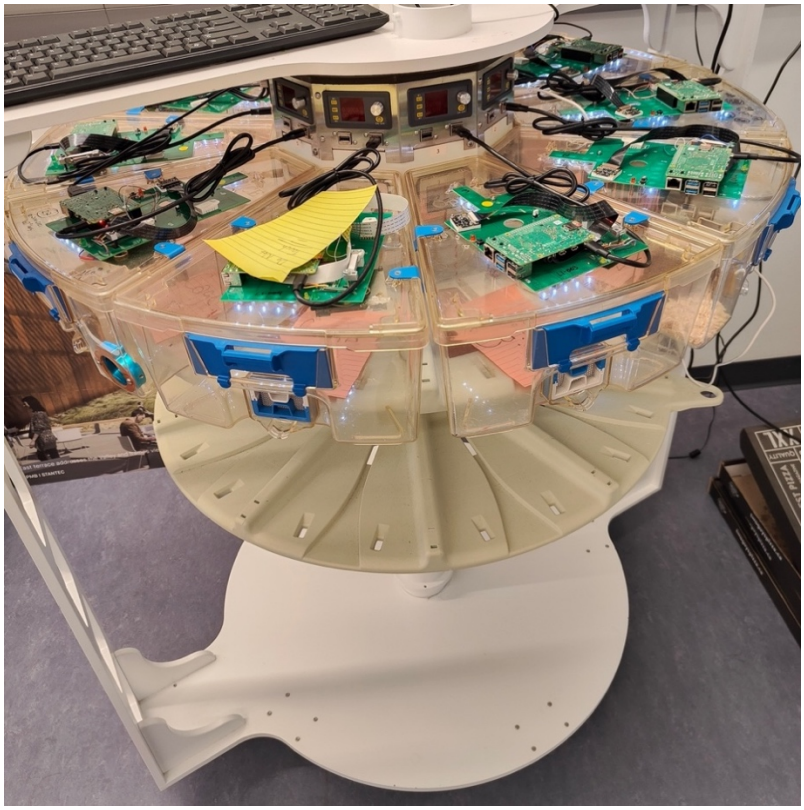

Figure 1: Homecages and rack. An example setup of the homecages. Separate sets were used in this thesis.

Each homecage top was integrated with a RaspberryPi (RPi) 3B+ computer, a camera module with fisheye lens on an overhead angle, LED lighting and other sensors (Singh, Bermudez-Contreras, Nazari, Sutherland, & Mohajerani, 2019). A homecage control application was developed to operate the homecage system and store recorded data. Captured data was stored on a separate partition in the MicroSD where the operating system and the homecage application are stored. Another data transfer application was developed and ran on a separate computer to collect data continuously from the RPi and to archive the data in local storage after some post-processing. Both the homecage control application on the RPi's and the data transfer application on the computer were written in Python and ran on Linux based operating systems (Raspberry Pi OS 9 and Ubuntu 20.04 respectively).

Procedure

Behavioural recording started 4 weeks after move-in and lasted 26 weeks. Once the recording started, it remained running continuously for the entire duration of the experiment. Video was recorded at a resolution of 640x480 pixels at a framerate of 10 frames per second (FPS) and encoded in MP4 format. Recordings were broken into 20 minutes segments for data transfer purposes, and recording was resynchronized at 12:00 am every day. The room housing the mice remained closed to ensure that there were no outside disturbances to the mice. No personnel entered the room outside of the occasions described below.

Animal care staff entered and inspected the mice daily. In these instances, animal care staff performed visual inspection through the transparent cage wall, and the cages were not moved or opened. Once a week, animal care staff performed full inspection and cage maintenance, usually on Tuesday and occasionally on Wednesday. This involved opening the cages, examining and weighing the mice, refilling or replacing the food hoppers and water bottles if needed. The cage bottoms were scheduled to be replaced every two weeks during cage maintenance. To avoid sudden change to clean cages causing a disturbance to the mice, each homecage was connected to a side cage with a tube one week prior to the change so the mice could explore and scent the cage. During the change, the entire cage bottoms of the homecages were replaced with those of the side cages. The food hoppers, water bottles and half of the nests were also transferred to the side cages. Additional nesting materials were added.

Two days each week (usually the day before and after weekly cage maintenance from animal care staff) were reserved for maintenance and experiment related tasks from experimenters. Tasks during these maintenance periods included hardware maintenance such as camera lens cleaning, and software maintenance such as system reboots, as well as other tasks such as setting up partitions and connecting tubes.

All human presence in the experiment room outside the daily inspection were logged, including additional unscheduled maintenance related to animal health and critical software/hardware error. This research was approved by the University of Lethbridge Animal Welfare Committee under protocol 1812-01 and 2209-01. It was conducted according to standards from the Canadian Council on Animal Care (CCAC).

Analysis

Video tracking technique was used on the video data recorded from the experiment for analyses of the circadian rhythm and excursion behaviours. This section will introduce the pipelines of the analyses.

Mouse tracking. Deep learning object detection model YOLOv8 (You Only Look Once) (Jocher, Chaurasia, & Qiu, 2023) was used to track the mice in the homecage videos. To train the YOLOv8 model, a total of 2283 frames were selected from all available videos. The selection process ensured frames from all cages at both light and dark cycles were included. The selection intentionally included a good portion of "difficult" images, such as background images with no mice, images where the mice were obstructed, and images where the mice were heavily distorted from fisheye lens. The data

set was then separated into a training set with 1528 frames, a validation set with 427 frames and a testing set with 328 frames. As the training set was relatively small according to YOLOv8 recommendation, the image augmentation library Albumentations (Buslaev, et al., 2020) was used to increase the diversity of the training set. Albumentations was integrated into YOLOv8 to apply random permutations including colour variation, blurriness, rotations, reflection, zooming, etc. to the images during the training process. Also, to account for the small training set, the model was trained from the large detection model yolov8l.pt supplied by YOLOv8 through transfer learning. The training went for 100 epochs, and the performance was tested against the validation set after every epoch. After the training stopped, the performance of the detection model was assessed through testing on the testing set. The choice of training for 100 epochs was based on comparisons between models trained for 100, 200 and 300 epochs. The models trained for 200 and 300 epochs showed worse performance due to overfitting.

Three consecutive days of video data were selected every two weeks for analysis for all cages. The data were taken from the weeks where the homecages were not connected to the side cages. Data from days with logged human presence and any other technical issues were excluded from the analyses. The detection model was then applied to these selected data to track the location of the mice in their homecages in each individual frame. The resulted x and y coordinates of the centre of the bounding box predicted by YOLOv8 were taken as the location of the mice. In the cases where the locations of the mice were missing, such as in the cases where the mice were obstructed, their last known locations were used as an approximation.

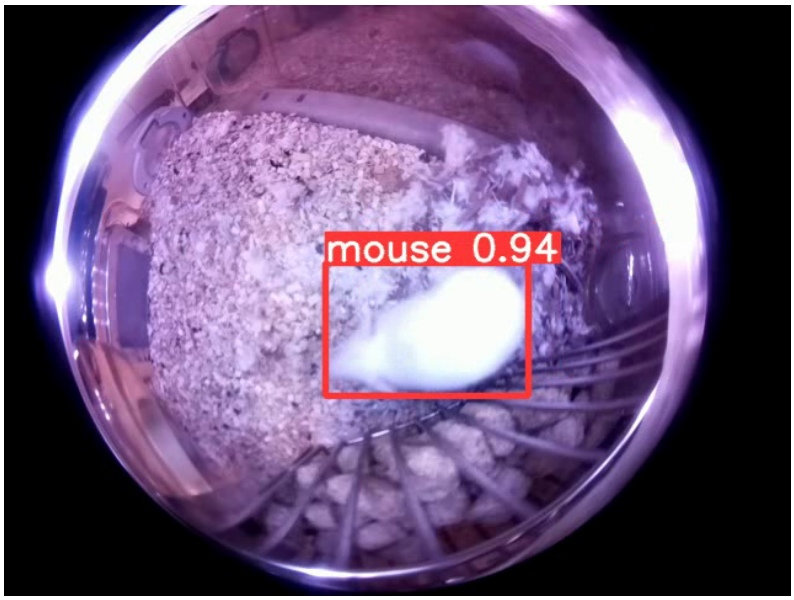

Figure 2: Homecage view and the bounding box. The top-down view of the homecage. The red bounding box and number are the mouse detection by YOLOv8 and its confidence.

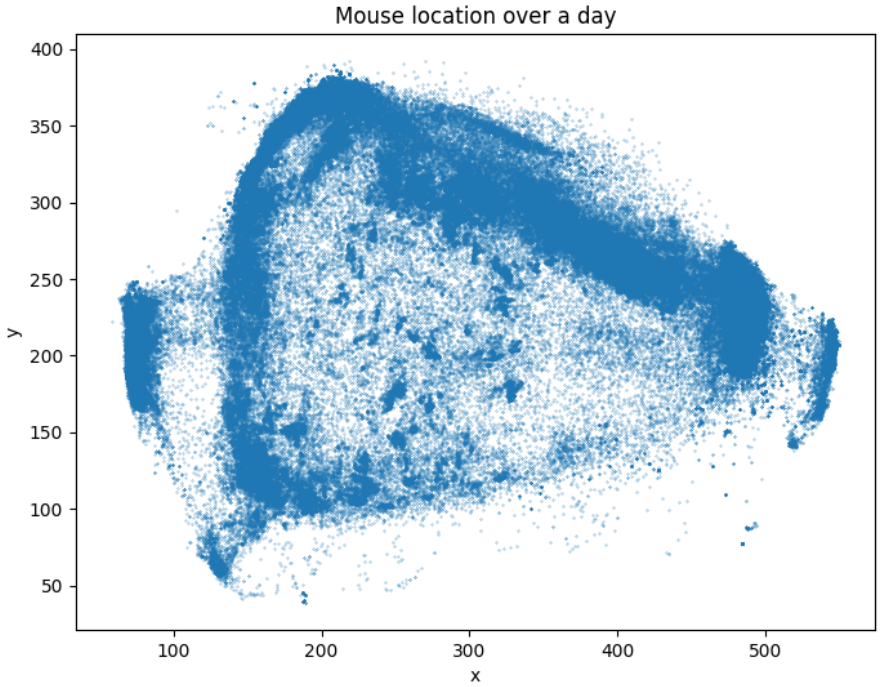

Figure 3: Locations of a mouse. The predicted location of the mouse in a cage over one day. Note that the Y-axis is reversed in this plot comparing to the view of the video.

Activity and Sleep. As videos were recorded at 10 FPS, averages and standard deviations of x and y coordinates every 10 frames were taken to represent the corresponding values for each second in time. With the standard deviations of x and y coordinates for every second, Pythagorean theorem was used to reduce the two standard deviations to one single value representing the spread of the mouse's location within that each second. A threshold of 1 pixel was applied to the spread each second. If the spread was above 1, the mouse was predicted to be active in that second, else the mice was predicted to be inactive. The threshold was chosen based on both visual comparison between the estimation and video, and comparison between analyses results and established activity pattern of mice such as total hours of sleep per day.

The active and inactive states each second were aggregated for every 60 seconds to provide an estimation of the number of seconds the mouse remaining active for every minute. This value was used as an estimation of the activity level for each minute. Average activity in the light cycle, dark cycle and full day were compared between the experiment and control groups throughout the experiment.

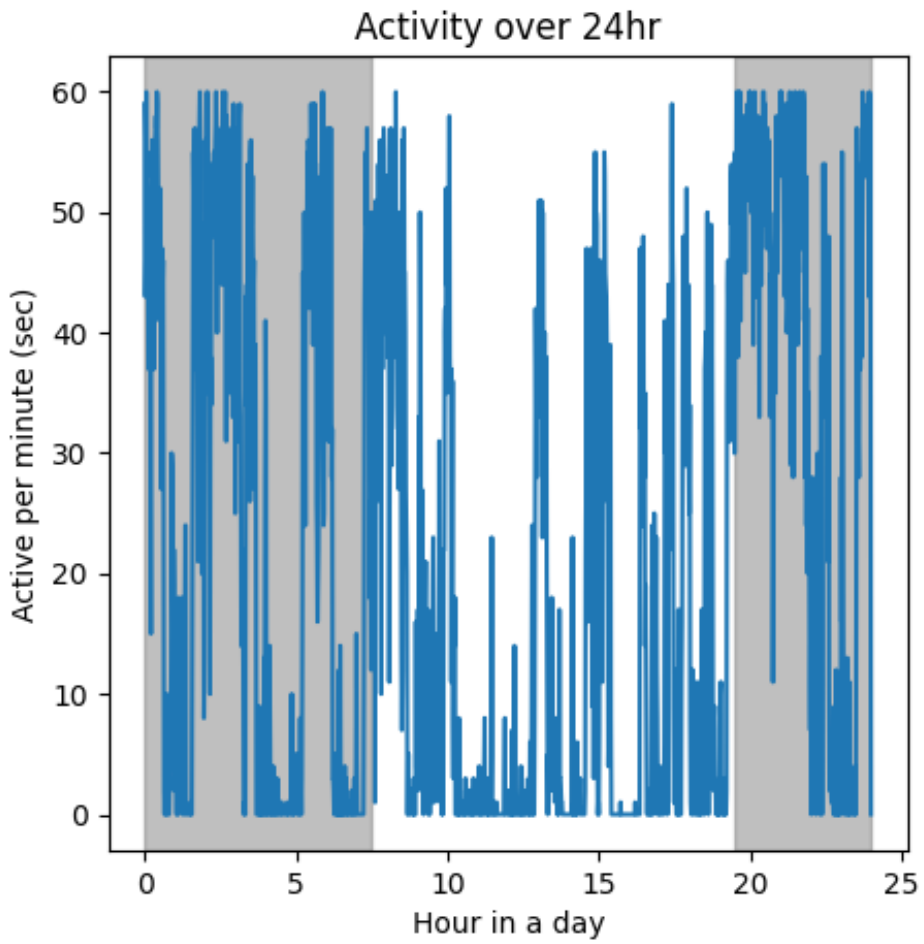

Figure 4: A mouse's activity level over 24 hours in the homecage. The shaded area is the time of the dark cycle.

The active and inactive states each second were also used to estimate the sleep states of the mice. When a mouse was in a continuous inactive state for over 40 seconds, the mouse was considered to be in the sleep state (Singh, Bermudez-Contreras, Nazari, Sutherland, & Mohajerani, 2019). Such continuous inactive periods were used as approximations to sleep bouts. The total duration of sleep, number of sleep bouts and average bout length were compared between the experiment and control groups throughout the experiment period. In addition, the histograms of sleep bout length were

generated each day for each mouse throughout the experiment. The histograms in the first 12 weeks and the last 12 weeks of the experiment were averaged to obtain the distribution of the sleep bout length in the respective experiment groups and time ranges. All these analyses were performed for data from full day, light cycle only and night cycle only.

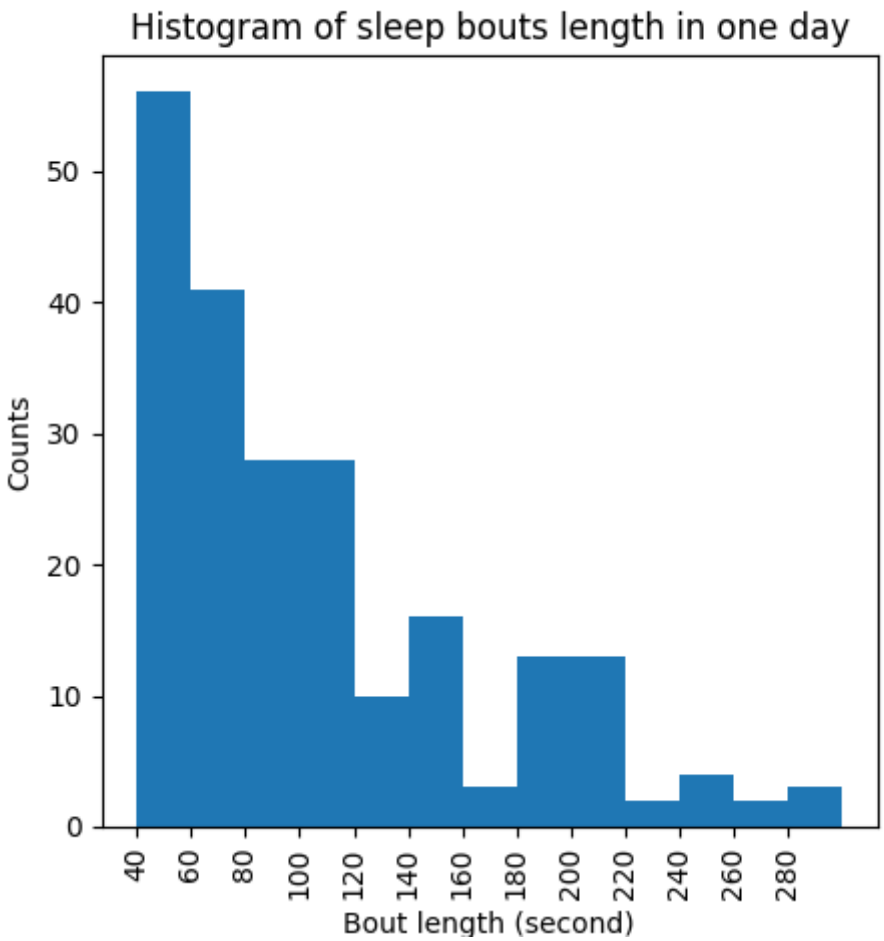

Figure 5: Histogram of sleep bouts. An example of a histogram of the length of the sleep bouts for one mouse over one day. Only sleep bouts of up to 300 seconds are displayed.

Periodicity. Instead of the per minute activity level, the running mean over a 15-minute window were used for periodicity analysis to reduce the effect of higher frequency oscillation. The daily activity level running means within the first 12 weeks and last 12

weeks among the same mouse group were concatenated together respectively. The Fourier transform was then applied to convert the data into frequency domain and the power spectral density was calculated (Duhamel & Vetterli, 1990). The power spectral density was normalized based on both the length of the data in time domain and the peak of the power spectral density for the purpose of comparison between the control and experiment groups. The power spectral diagrams were compared between the two groups to investigate the effects of AD progression to the circadian rhythm (in the lower frequency range) and basic rest-activity cycle (BRAC - in the higher frequency range).

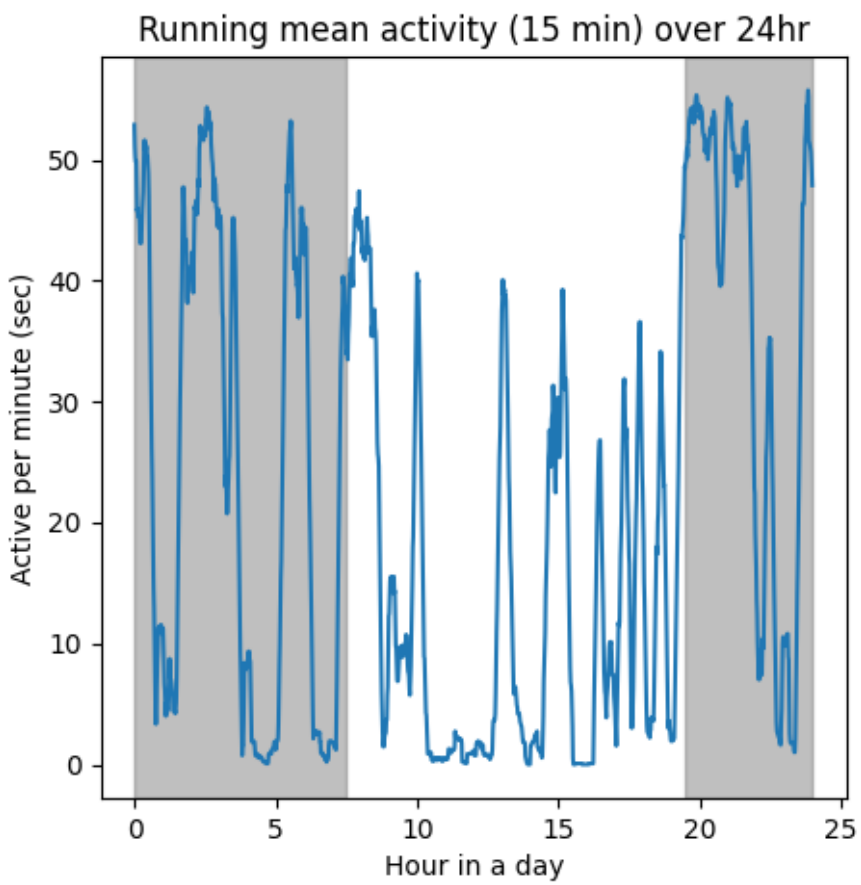

Figure 6: The 15-minute running mean of activity level over 24 hours. The shaded area represents the dark cycle.

Excursion behaviours. The average positions per second of the mice over a full day were used in the excursion analyses. First, kernel density estimation with Gaussian kernel was used to detect the coordinates where points (average positions) were most densely populated. That was the area where the mice spent most of the time over a day. As a result, the area within a 100-pixel radius of the coordinates was defined as the homebase (Bermudez-Contreras, Sutherland, Mohajerani, & Whishaw, 2022). This area generally coincided with the nests of the mice, and the selection of 100-pixel also took into consideration of the size of the nest in the video. With the homebase defined, one excursion was defined as the path of the mice travelled from them leaving the homebase area to them returning to the homebase area. More specifically, it was based on the mice estimated positions leaving and returning to the defined homebase area. Only excursions of at least 3 seconds and at most 60 seconds in duration were to be considered. Longer excursions than 60 seconds while rare can happen, but they are also at risk of being the results of mislabeling. In addition, a stop was defined as each time the mice stayed in the same position for over a second. The average excursion duration, average excursion distance, and average number of stops per excursion in light cycle, dark cycle and full day were compared between the experiment and control groups over time.

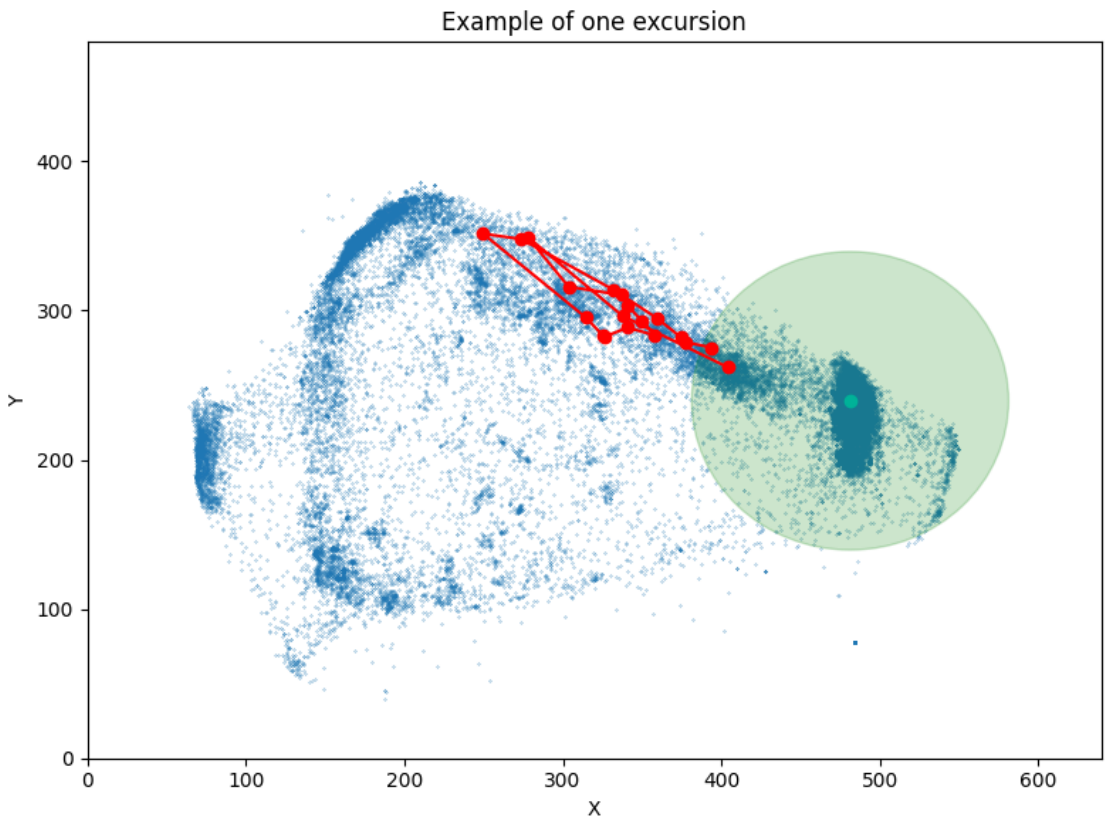

Figure 7: Example of one excursion. An example of one typical excursion. The blue points are the average position of the mouse over every second of a day. The large green dot is the estimated position where the activities of the mouse are most densely populated. The green shaded area is the estimated homebase area. The red line is the path of the excursion.

The histograms of excursion durations, excursion distances and number of stops were generated each day for each mouse throughout the experiment. The histograms in the first 12 weeks and the last 12 weeks of the experiment were averaged to obtain the distributions of the excursion durations, excursion distances and number of stops in the respective experiment groups and time ranges. The distributions were separately compared between the two groups and the two time ranges. In addition, the numbers of excursions for each hour in a day were also counted for all mice and averaged over the

two groups and the two time ranges (the first 12 weeks and the last 12 weeks of the experiment) for comparison.

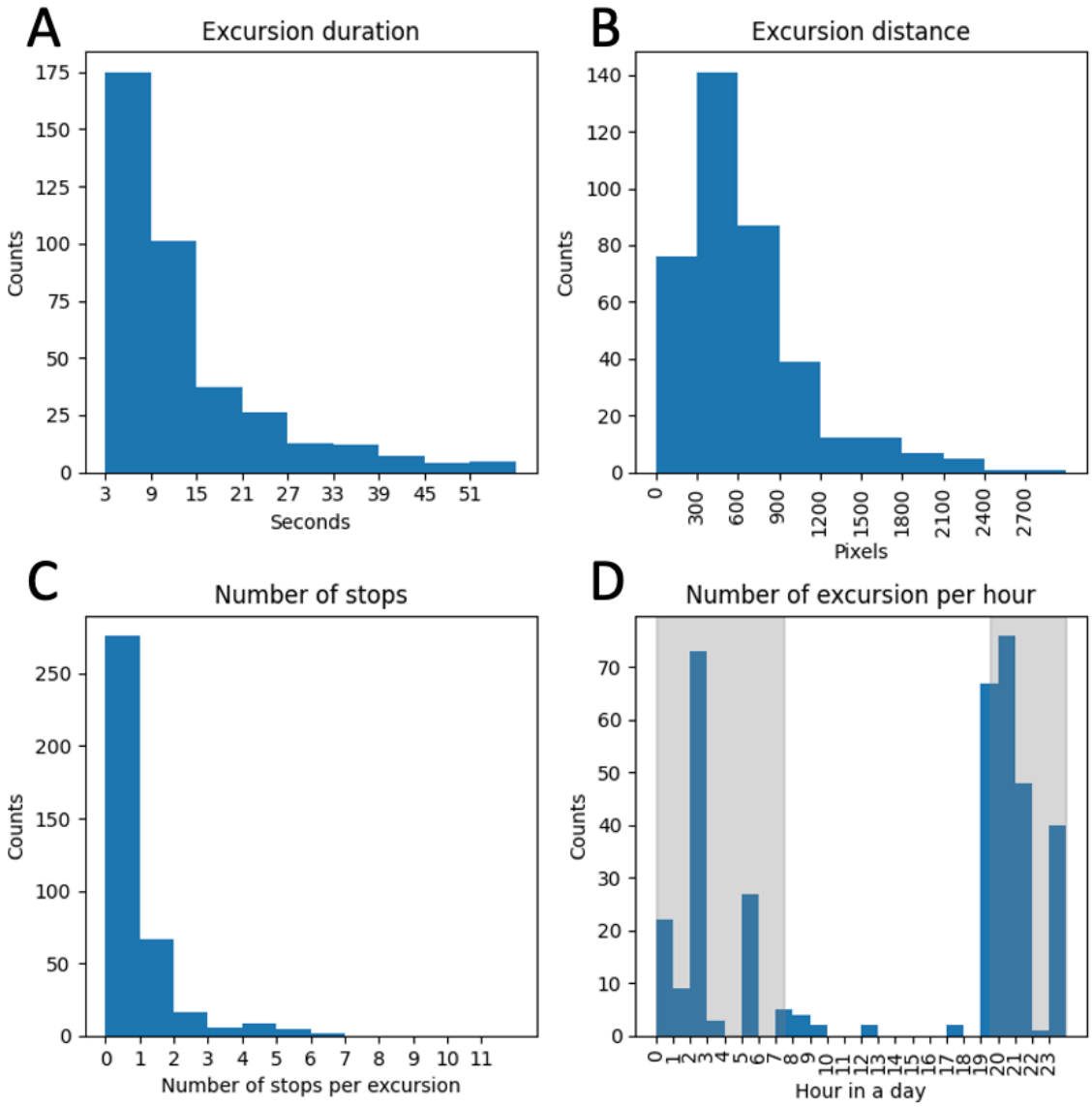

Figure 8: Excursion behaviour for one mouse in one day. Histograms of excursion behaviour of one mouse over one day. A: histogram based on the duration of the excursion. B: histogram based on the distance of the excursion. Only excursions with distances up to 3000 pixels are shown. C: histogram based on the number of times the mouse stopped for more than 1 seconds. D: numbers of excursion for each hour within a day, shaded areas are the dark cycle.

Statistical Analysis

For all comparisons between the experiment and control groups over time, two-way mix-model design ANOVA tests, with age being the within-subjects variable and experiment groups being the between-subjects variable, were first used to test the differences between the two groups. After differences were detected, pair-wise Mann-Whitney U tests were done to compare the data from the two groups in the same weeks.

Result

Mouse Tracking Model

To evaluate the trained YOLOv8 mouse tracking model, common performance metrics including precision, recall and mean average precision (mAP) were obtained from YOLOv8's results (Handelman, et al., 2019). Precision measures how accurate is the model's positive predictions. It is defined as the fraction of true positive predictions among all positive predictions, which include true positive and false positive. In the context of the experiment, true positive were the cases where the model correctly identified the mice, and false positive were the cases where the model misidentified other objects as the mice. Recall measures how well the model identifies all positive cases within the data. It is defined as the fraction of true positive predictions among all positive cases, which include true positive and false negative. In the context of the experiment, false negatives were the cases where the model failed to identify the mice. By varying the confidence threshold, a precision-recall (PR) curve can be made, and mAP is defined as the area under the PR curve.

Note that for object recognition using bounding box, including YOLOv8, the evaluation between true positives and false positives are based on the metric intersection over union (IoU), which is measured by the ratio of the area of intersection, between the predicted bounding box and the group truth, and the area of the union between the two. If the ratio is above a defined IoU threshold, then it is considered true positives, otherwise it is considered false positive. This leads to two more specified performance metrics involved mAP. One is mAP@0.5, which is mAP at IoU threshold of 0.5. Another is

mAP@0.5:0.95, which is the average of mAP based on IoU ranged from 0.5 to 0.95 with an increment of 0.05.

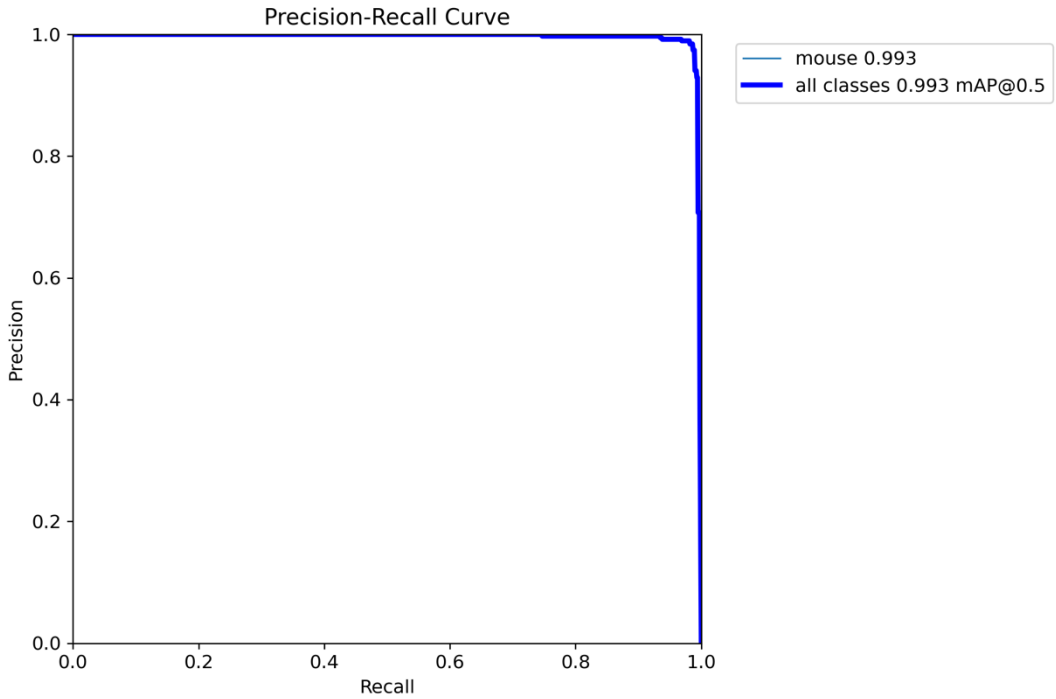

Figure 9: Precision-Recall curve. The PR curve of the YOLOv8 mouse tracking model showed the extremely high performance of tracking.

The trained YOLOv8 mouse tracking model had a precision of 0.983 and a recall of 0.987, based on YOLOv8's default confidence threshold of 0.25. The mAP@0.5 was 0.993, and the mAP@0.5:0.95 was 0.793. These were very high scores in performance metrics that might imply overfitting. However, since the model was not intended to be generalized to data outside of this experiment, the high score is acceptable. Manual verifications with randomly selected video were done to further ensure prediction accuracy.

Circadian Rhythm

Activity level. Both groups started with similar levels of daily average activity at the beginning of the experiment. Throughout the experiment, the control group maintained a relatively stable daily activity level with a small downward trend, which is common as mice age (Pernold, Rullman, & Ulfhake, 2021). The daily activity level of the experiment group, however, followed an upward trend. Mixed model ANOVA test results showed a significant interaction between age and groups ($F[12,48] = 3.494, p = 0.000953, \eta_p^2 = 0.466$). Pair-wise Mann-Whitney U tests showed that the activity level between the two groups started to differ significantly at the age of 25 weeks and became consistently different significantly ($p < 0.05$) starting from week 33. Mixed model ANOVA test results for the average activity level in the light cycle also showed a significant interaction between age and groups ($F[12,48] = 2.216, p = 0.0258, \eta_p^2 = 0.357$). However, pair-wise Mann-Whitney U tests did not show statistically significant difference until the last week of the experiment. The dark cycle activity level follows the same trend as the full day level. Mixed model ANOVA test results for the average activity level in the dark cycle also showed a significant interaction between age and groups ($F[12,48] = 2.733, p = 0.00670, \eta_p^2 = 0.406$). Pair-wise Mann-Whitney U tests showed significant divergence started at week 25 and the activity levels became consistently different significantly from week 33. This showed that the change in activity level predominately happens in the dark cycle during the dark cycle.

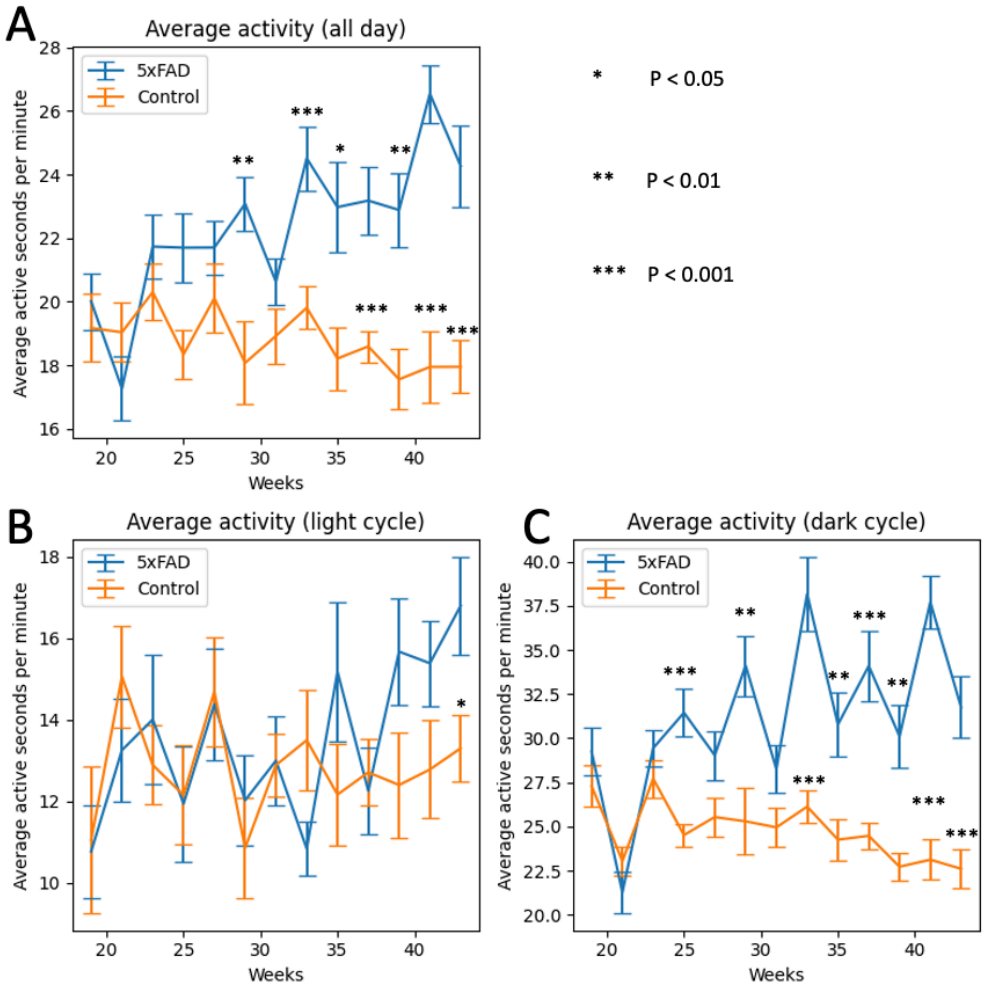

Figure 10: Comparison of average activity. Comparing full day (A), light cycle (B) and dark cycle (C) average activity with all cages and all days between the two groups. The error bars are standard error, and the statistical tests used are Mann-Whitney U tests. The average activity diverged between the two groups, and it happened mainly in the dark cycle.

Sleep. The amount of sleep followed a opposite trend as the activity level. At the beginning of the experiment, both groups spent about 50% of time sleeping, which is common for lab mice (Campbell & Tobler, 1984). Throughout the experiment, the control group maintained a stable amount of sleep around 50%. The experiment group on the other hand dropped below 45% from week 23, and only slept on average 35% of the time in a day at the end of the experiment. Mixed model ANOVA tests showed significant

interaction between age and groups for all three cases of full day ($F[12,48] = 3.531, p = 0.00087, \eta_p^2 = 0.469$), light cycle ($F[12,48] = 2.245, p = 0.023915, \eta_p^2 = 0.360$) and dark cycle ($F[12,48] = 3.189, p = 0.00207, \eta_p^2 = 0.444$). Similar to what was observed for the activity level, the amount of sleep in the light cycle did not differ significantly between the two groups until the last month of the experiment, while the divergence happened much earlier in the dark cycle.

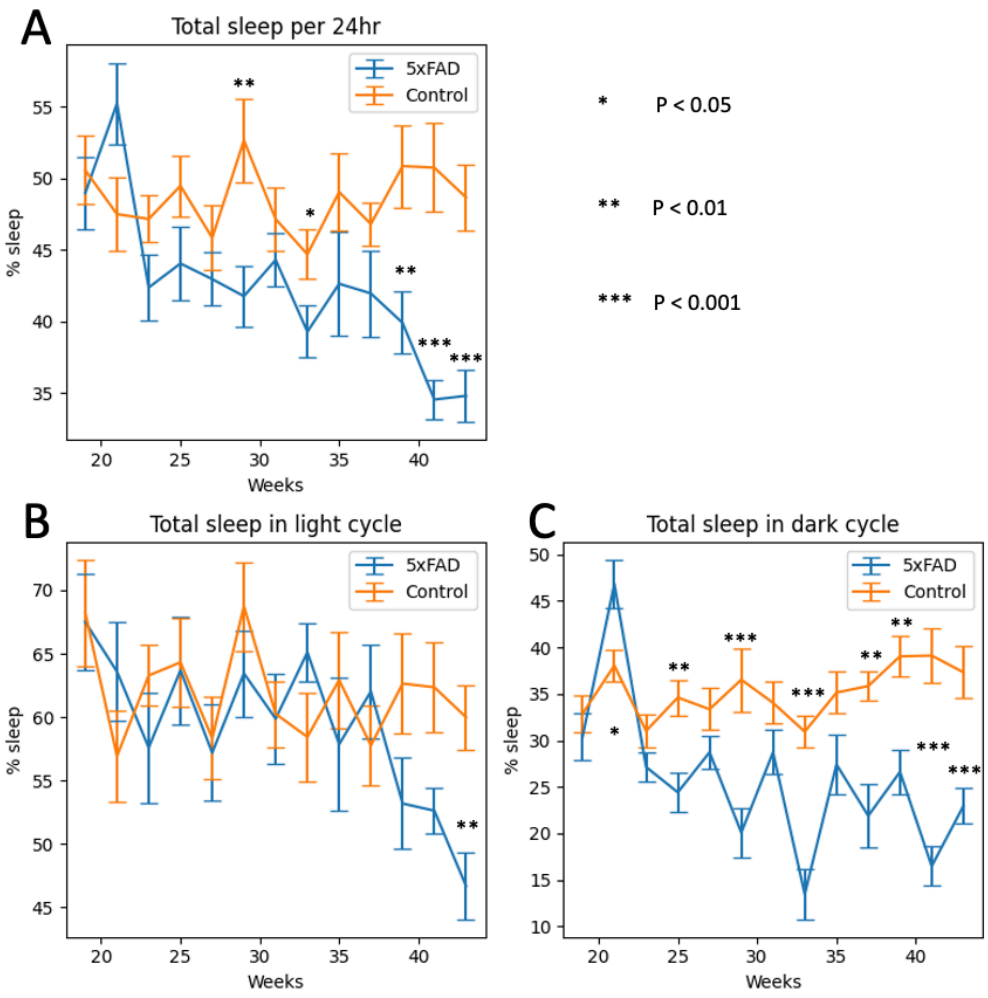

Figure 11: Comparison of the amount of sleep. Comparing the average amount of sleep as a percentage of the entire period of a day (A), light cycle (B) and dark cycle (C) with all mice and all days between the two groups. The error bars are standard error, and the statistical tests are Mann-Whitney U tests. Non-transgenic mice maintained a stable amount of sleep time, while 5xFAD mice's sleep time dropped over time.

However, sleep bout analyses showed no clear differences between the two groups throughout the experiment in full day, light cycle, or dark cycle. For average number of sleep bouts, mixed model ANOVA tests showed no significant interaction between age and groups for all three cases of full day ($F[12,48] = 0.359, p = 0.972, \eta_p^2 = 0.0824$), light cycle ($F[12,48] = 0.414, p = 0.951, \eta_p^2 = 0.0937$) and dark cycle ($F[12,48] = 1.104, p = 0.379, \eta_p^2 = 0.216$). For average duration of sleep bouts, mixed model ANOVA tests also did not showed significant interaction between age and groups for all three cases of full day ($F[12,48] = 0.810, p = 0.639, \eta_p^2 = 0.168$), light cycle ($F[12,48] = 0.471, p = 0.922, \eta_p^2 = 0.105$) and dark cycle ($F[12,48] = 1.180, p = 0.324, \eta_p^2 = 0.228$).

In addition, the comparisons of the distribution of sleep bout lengths between two groups (experiment and control) and between two time ranges (first 12 weeks and last 12 weeks) showed that non-transgenic mice generally had more bouts of sleep across the ranges of sleep bout lengths being compared (40 to 300 seconds). The only noticeable difference between the two time ranges was in the light cycle. Both groups showed an increase in sleep bouts across the different lengths, but the experiment group showed a bigger increase, especially in the lower ranges of the bout lengths.

.

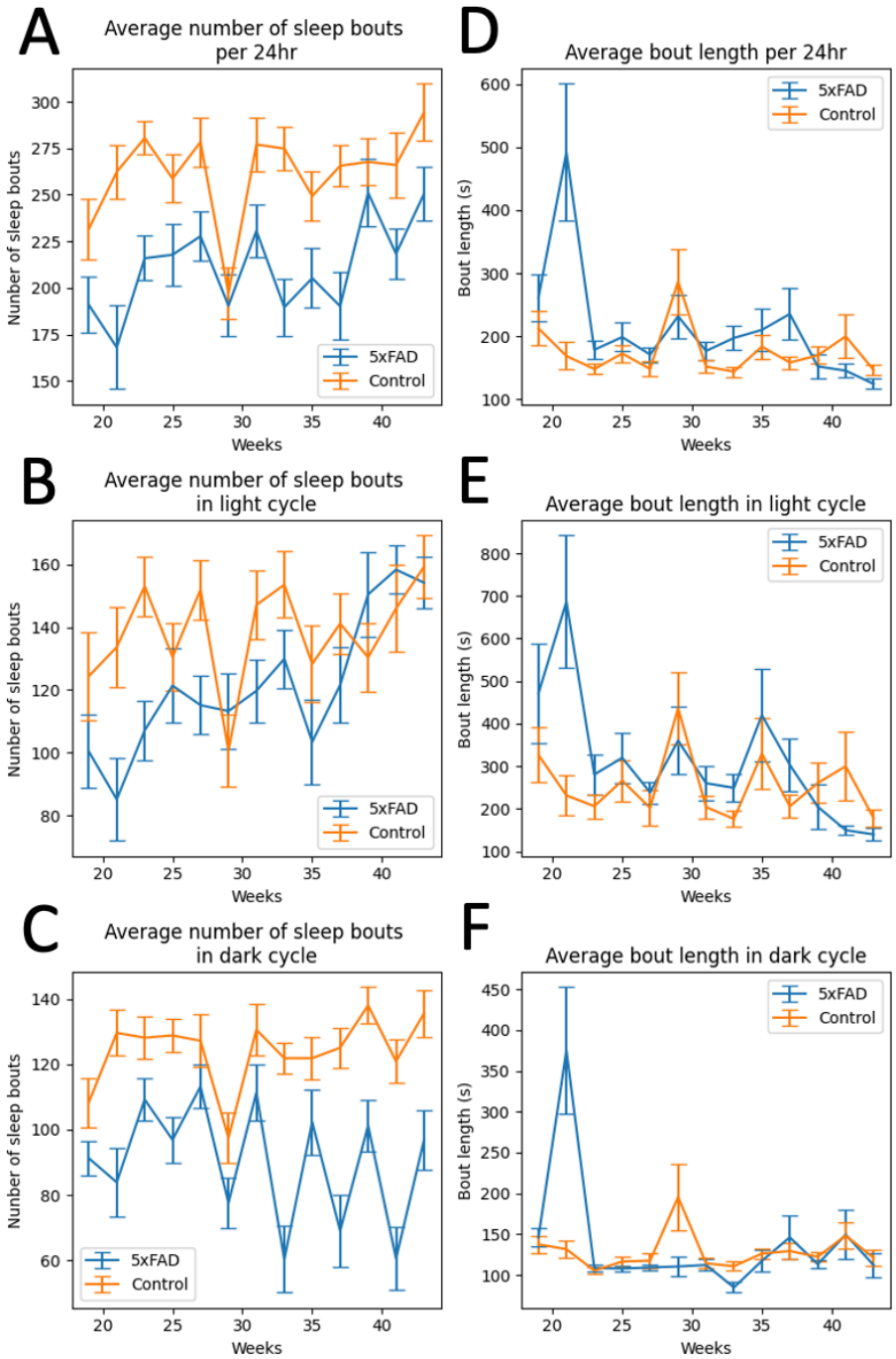

Figure 12: Comparison of sleep bouts over time. Comparing the average number of sleep bouts of the entire period of full day (A), light cycle (B) and dark cycle (C) with all mice and all days between the two groups. Also comparing the average length of sleep bouts in full day (D), light cycle (E) and dark cycle (F). The error bars are standard error. 5xFAD mice had lower number of sleep bouts throughout the experiment, but it was not statistically significant according to mixed model ANOVA. Both groups have similar sleep bout length over time. The high bout length of 5xFAD mice in the second point was likely an outlier resulting from tracking issue.

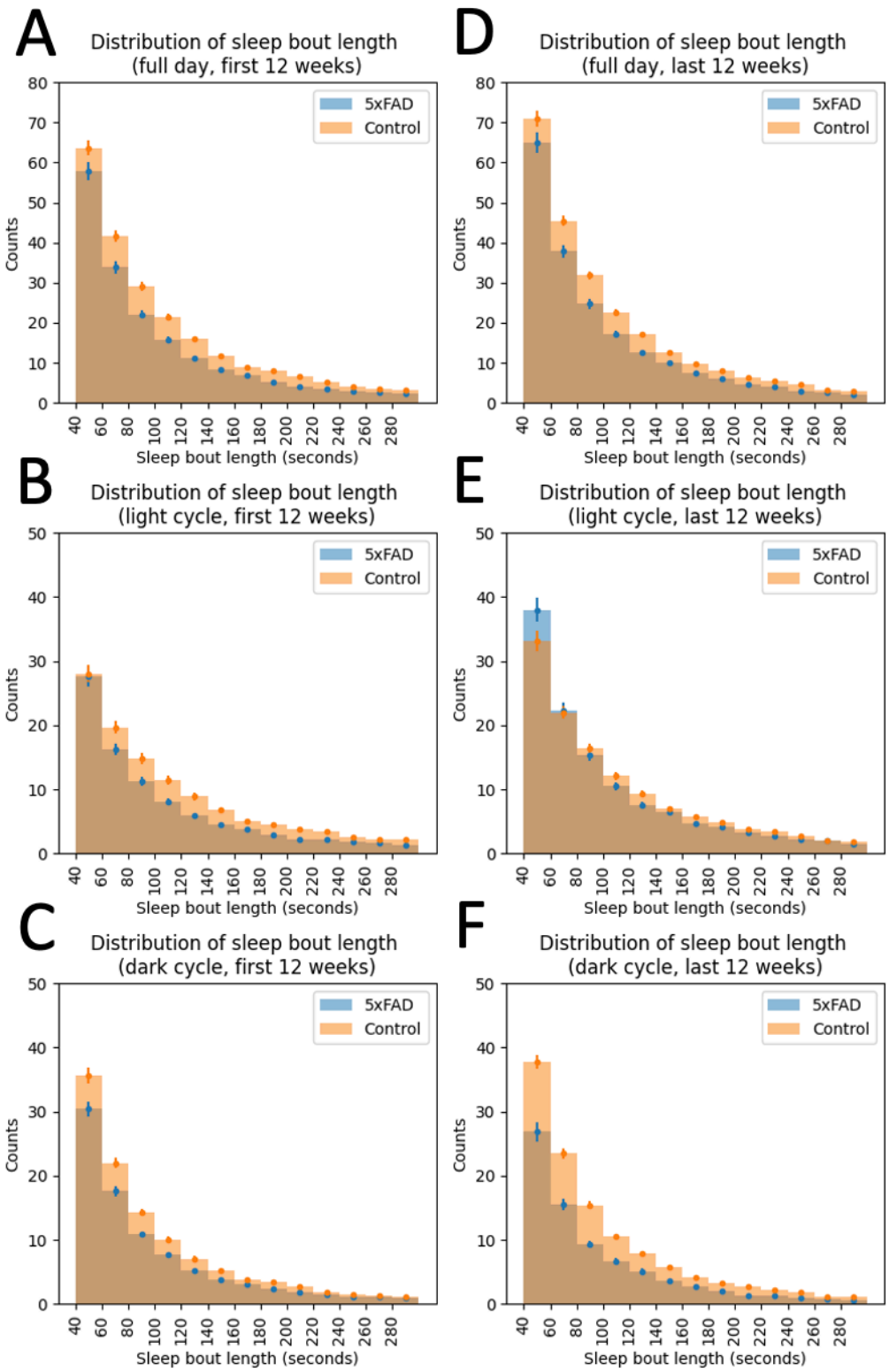

Figure 13: Comparison of the distributions of sleep bout lengths. Comparing the averaged sleep bout lengths distribution between the two experiment groups and two age periods. The error bars are standard error. One notable difference in the distribution was the greater increase in shorter sleep bout length in the light cycle of the last 12 weeks.

Periodicity. Fourier analysis in lower frequency range showed that the experiment group had slightly higher normalized power spectral density for the activity level at a period of 24 hours, which corresponded to the circadian rhythm. This happened in both the first 12 weeks and the last 12 weeks of the experiment. In addition, peaks at 12-hour, 8-hour, 6-hour, and 4-hour periods were observed in both time ranges. In shorter period range, both the experiment and control groups exhibited a comparable level of normalized power spectral density in the first 12 weeks of the experiment. However, in the last 12 weeks of the experiment, non-transgenic mice had overall higher level of normalized power spectral density than 5xFAD mouse models, especially in the sub-hour period range. These results showed that the circadian rhythm of the 5xFAD mouse model was not affected by the disease progression, but the BRAC in the sub-hour range were impacted.

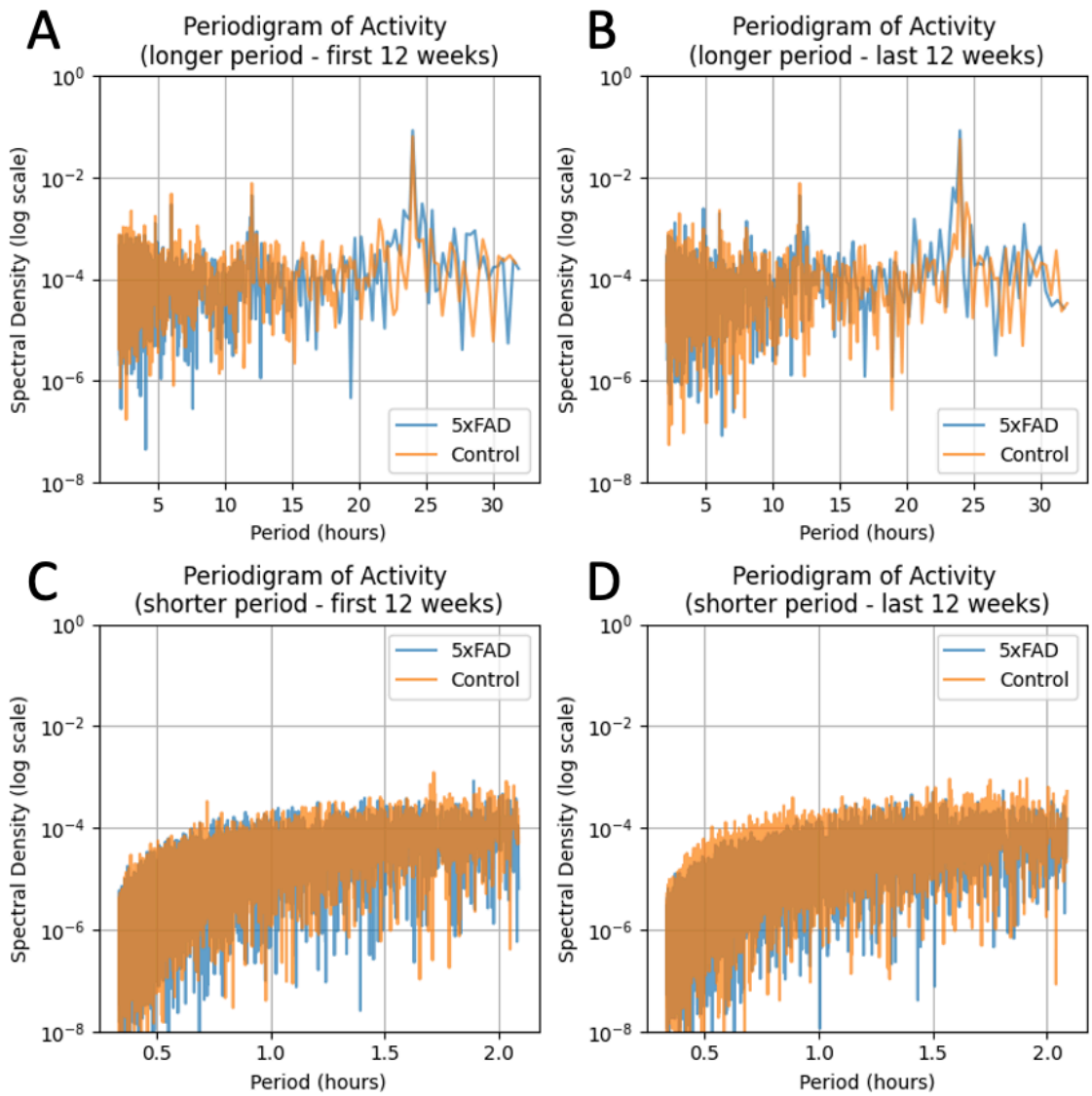

Figure 14: Periodograms of activity. Comparison of power spectral density of activity level for the first 12 weeks and the last 12 weeks, in log scale. Data sample frequency is 0.01667 Hz. Activity data from all mice and all day in the same group and period are concatenated for the analysis. There was little change in longer period range. The 5xFAD mice had lower spectral density in shorter period range in the last 12 weeks, especially in the sub-hour period range.

Excursion Behaviour

The experiment group had a higher total number of excursions per day than the control group throughout the entire experiment period. Similar to the activity level, the number of excursions followed a slight upward trajectory for the experiment group, while the control group had a slight downward trend. In majority of the experiment period, the control group had a higher average duration and number of the stops for the excursions. 5xFAD mice had an increase trend in the average duration of excursion, but a decrease trend in average distance of excursion. This showed a decrease in average speed of the movement during excursion. However, the two groups did not show significant interaction between groups and ages in all four of total number of excursions ($F[12,48] = 0.621, p = 0.814, \eta_p^2 = 0.134$), average duration of excursions ($F[12,48] = 1.542, p = 0.142, \eta_p^2 = 0.278$), average distance of excursions ($F[12,48] = 1.038, p = 0.431, \eta_p^2 = 0.206$) and average number of stops ($F[12,48] = 0.895, p = 0.558, \eta_p^2 = 0.183$).

The distributions of excursion duration, distance, and number of stops all displayed a similar trend. The non-transgenic mouse models showed a decrease across the lower end of the distribution in all three categories in the last 12 weeks comparing to the first 12 weeks. The 5xFAD mouse models showed a big increase in the number of shorter distance excursions, which explained the decrease in average excursion length. The 5xFAD mouse models also showed no increase to small decrease in number of shorter duration excursion and number of excursions with fewer stops. Also worth noting that in the first 12 weeks, the 5xFAD model already had higher level in all three distributions.

This might be caused by the fact that in the first 12 weeks of the experiment, the mice already reached the age of 29 weeks, at which point 5xFAD mice might already display behavioural effects from disease progression.

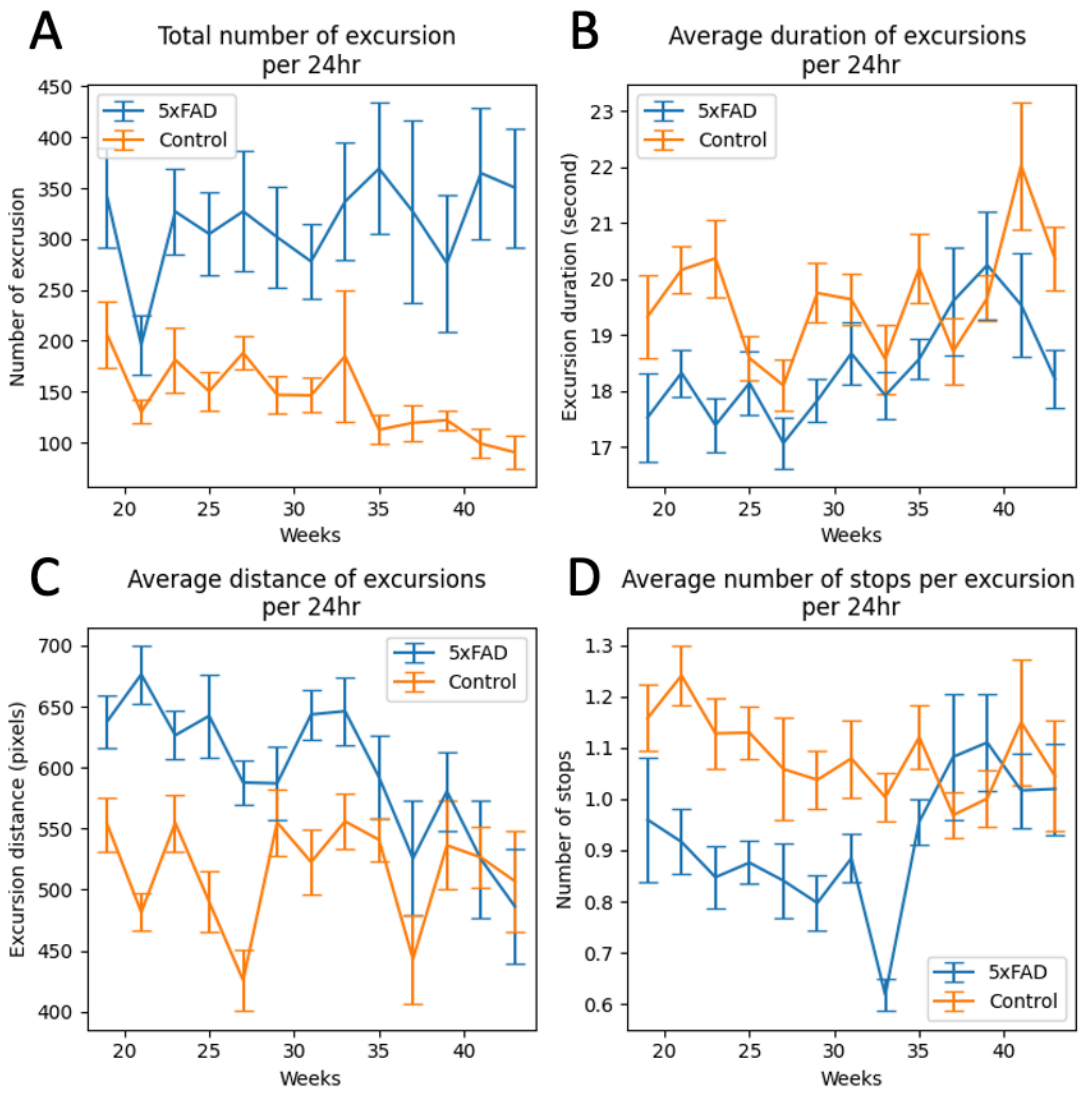

Figure 15: Comparison of excursion behaviours. Comparing total number of excursion (A), average duration of excursion (B), average distance of excursion (C) and average number of stops per excursion (D) per 24 hours between the two groups. The error bars are standard error. 5xFAD had higher number of excursions per day throughout the experiment. However, according to mixed model ANOVA, there is not enough statistical significance.

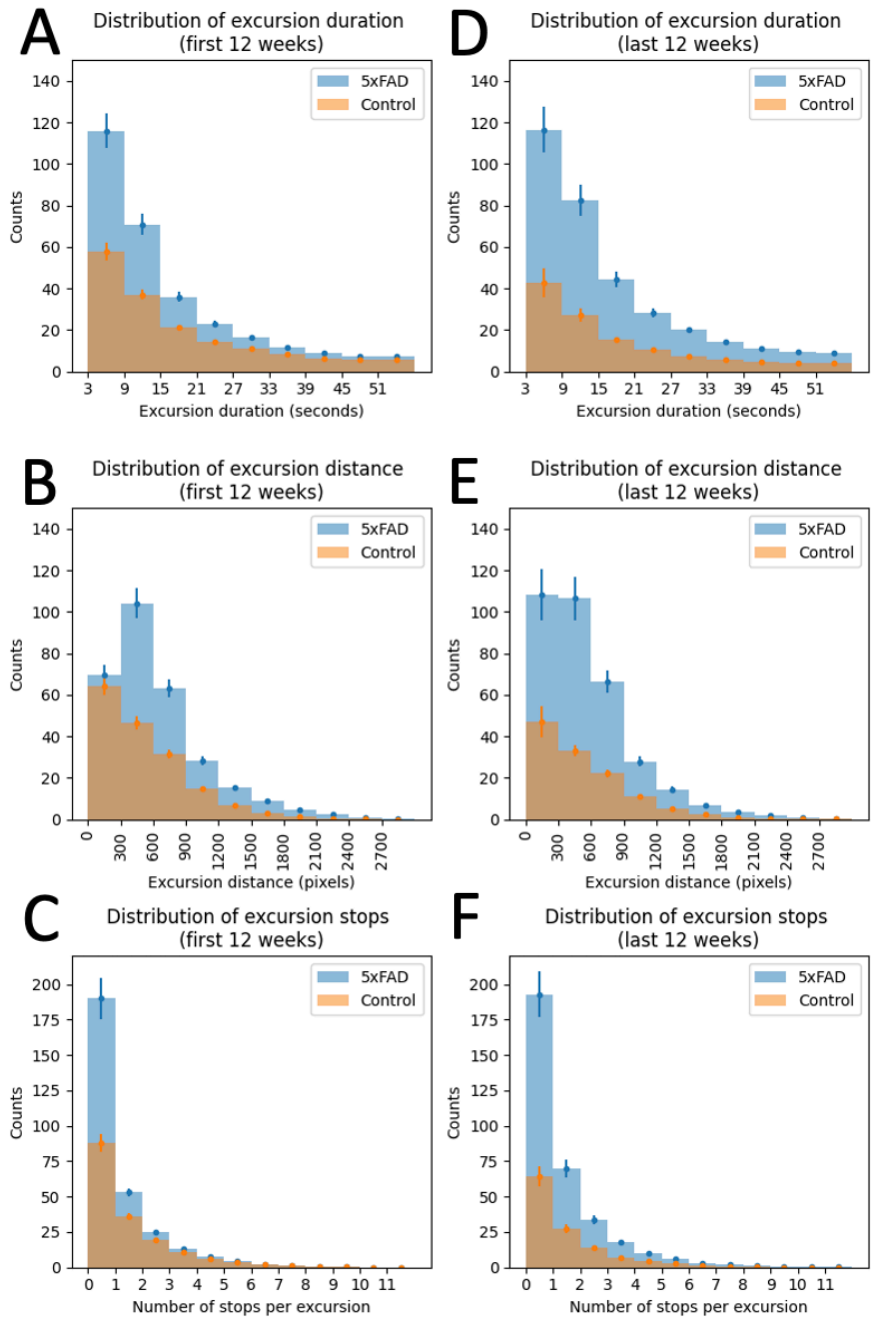

Figure 16: Comparisons of the distribution for excursion. Comparing the average distributions of excursion duration, distance and number of stops over both mouse groups and the two time ranges. The error bars are standard error. 5xFAD showed an increase in the number of shorter excursions in the last 12 weeks while non-transgenic showed no increase to small decrease in the number of shorter excursions.

Similarly, average numbers of excursions each hour over a day showed a higher number of excursions for 5xFAD mice in the first 12 weeks of the experiment comparing

to non-transgenic mice. This comparison was consistent with the activity level comparison. The changes in the numbers of excursions happened predominately in the dark cycle. Non-transgenic mice showed a decrease in the numbers of excursions in the last 12 weeks comparing to the first 12 weeks. When comparing the last 12 weeks with the first 12 weeks, the 5xFAD showed an increase in the numbers of excursions early in the dark cycle (second shaded area in the plot) and maintained a similar level late in the dark cycle (first shaded area in the plot). Another noticeable change in the last 12 weeks for 5xFAD mice was the more gradual ramp-up and ramp-down period over several hours before the transitions of the light/dark and dark/light cycle. This was not apparent for the 5xFAD mouse models in the first 12 weeks, or for the non-transgenic mouse models in both periods.

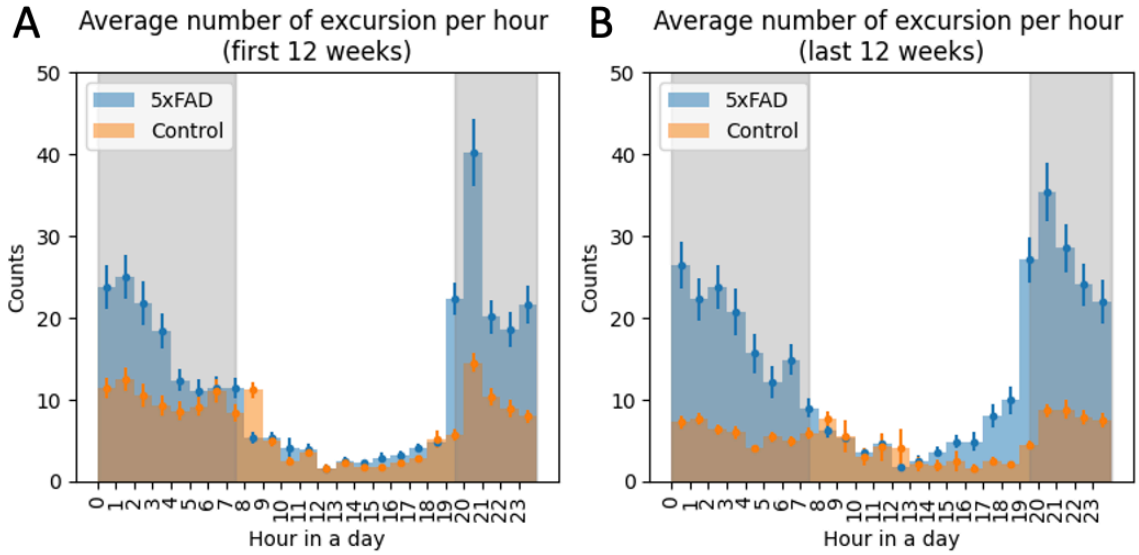

Figure 17: Comparison of numbers of excursions over a day. Comparing the average numbers of excursion for each hour over a day. The shaded area is corresponding to the dark cycle, and the error bars are standard error. Non-transgenic mice showed decrease in number of excursions per hour over a day in the last 12 weeks, while 5xFAD maintained at similar level. 5xFAD mice had more gradual ramp-ups and ramp-downs before the transition of light/dark and dark/light cycle in the last 12 weeks.

Discussion

This thesis proposed using homecage behaviours as a way to conduct long term studies of AD mouse models. The thesis aimed to detect differences in activity level, circadian rhythm and excursion behaviours over time between AD and non-AD mouse models. To accomplish these goals, a software application for controlling the automated homecage systems (Singh, Bermudez-Contreras, Nazari, Sutherland, & Mohajerani, 2019) and acquiring and pre-processing data was developed. In addition, a central server system, including all the corresponding software applications, dedicated for data collection, processing, visualization and archival was designed and developed. A 26-week long experiment with six female 5xFAD mouse models and six of their female non-transgenic littermates was developed and conducted for long term recording of their homecage behaviours. To track the mice in their homecages, a YOLOv8 object recognition model was trained using the frames collected and labeled from the homecage recording. With the tracking data of the mice, the activity level, circadian rhythm behaviours and exclusion behaviour were then analyzed.

The major findings in the thesis include that as the mice aged, the activity level of the non-transgenic mice was relatively stable with a small downwards trend commonly associated with aging, but the activity level of the 5xFAD mice diverged at around 25 weeks old and followed an upward trend throughout the experiment. This divergence of 5xFAD mice could almost entirely attributed to the change of activity level in the dark cycle, while in the light cycle, there was no significant differences between the two groups until the last week of the experiment.

The sleep analyses showed that the non-transgenic mice spent a stable 50% of time sleeping each day, while the sleep time of 5xFAD mice continued to decline and reached around 35% of time in a day. Again, this trend came almost entirely from the dark cycle. There was no significant difference for the sleep bout length and number of sleep bouts between the groups. There was a larger increase in number of shorter sleep bouts in the light cycle in the last 12 weeks of the experiment comparing to the first 12 weeks. As the total amount of sleep in the light cycle maintained at a similar level in the last 12 weeks for the 5xFAD mice, this suggests the mice were having more fragmented sleep, and it agrees with other sleep studies of 5xFAD mice (Sethi, et al., 2015). Studies also show that from 6-month of age, 5xFAD mice start having decreased NREM (Non-rapid eye movement) sleep (Drew, Park, & Kim., 2023). Note that, as periods of motionlessness were used as approximations to sleep bouts, increased amount of small movement during sleep might also lead to increased amount of sleep bouts detected.

Fourier analysis did not detect much difference between the two groups and the two time ranges (first 12 weeks and last 12 weeks of the experiment) for the lower frequency range associated with circadian rhythm. However, in the last 12 weeks of the experiment, the non-transgenic mouse model had higher normalized power spectral density in the higher frequency ranges corresponding to the sub-hour cycles. This signaled a less robust BRAC for 5xFAD mouse models after 33 weeks of age. As in the same period sleep disturbance was observed for 5xFAD mice, this also supports the BRAC hypothesis which suggests BRAC is associated with the change of sleep stages in the sleep cycles (Kleitman, 1982).

Excursion analyses showed a downward trend in the total number of excursions per day for the non-transgenic mice, while the 5xFAD mice displayed a small upward trend. The 5xFAD mice also showed a decrease of average movement speed during excursions, which suggests worsening mobility due to disease progression. However, no significant difference was detected in the average duration and distance of the excursions and the average number of stops within each excursion. The average activity per hour over a day showed long ram-ups and ram-downs of activity level before the transitions of light/dark and dark/light cycle for 5xFAD mice in the last 12 weeks of the experiment. This suggests an increase in variance in activity onsets and offsets at the change of day/night cycle, which is a sign of circadian disruption (Brown, Fisk, Pothecary, & Peirson, 2019).

Also worth mentioning, all the applications developed during the thesis are not only applicable to experiments investigating the homecage behaviours, but also capable to serve as general monitoring system for animal husbandry purpose. The mouse tracking model, while not designed to be used as a general tracking system outside of the experiment, has such potential when trained with more labeled frames and rigorous testing.

Animals

The experiment used 5xFAD transgenic mouse model as experiment group. The 5xFAD mouse model displays early and aggressive progression of amyloid pathology, and the behavioural effects were observable at as early as 4 to 5 months old in specific experiments (Oakley, et al., 2006; Devi & Ohno, 2010; Kimura & Ohno, 2009). The

increase in activity level measured in the experiment agrees with some reports of 5xFAD mice being hyperactive (Flanigan, Xue, Rao, Dhanushkodi, & McDonald, 2014). Hyperactivity is also considered an NPS for dementia in human patients (van der Linde, Dening, Matthews, & Brayne, 2014). Another reason for the increase in measured activity level for 5xFAD mice might be the fact that they can develop seizures as they age. At least three of the six 5xFAD mice in the experiment group had shown some level of seizures in the later part of the experiment, and one of those three had to be euthanized before the end of the experiment. This experiment can be repeated with other AD mouse models, such as APP-NL-GF, which also start to develop symptom at roughly 6 months old but do not develop seizures.

Homecage System

The automated homecage system has several advantages comparing to behavioural tasks. Firstly, it reduces the workload on the experimenters significantly while also minimized the behavioural effects to the mice from experimenters' presence and handling of the mice. This both allows larger scale experiments to be conducted and reduces the difficulty in reproducing the experiments in separate locations. It allows observation over much longer period than normal behavioural tasks can cover, and it also allows observation of the animals' night cycle behaviours, which is difficult to investigate through behavioural tasks. This experiment was able to show that the changes in the behaviours of 5xFAD mouse models happened first in the dark cycle, and later escalated to the light cycle. Due to the large amount of data that can be recorded throughout the experiment, this can reduce the number of animals used in experiments. Furthermore, the data recorded from the automated homecage can be used for many analyses in addition to

the tracking data investigated in this thesis, such as food and water consumption, and defecation. This also allows fewer experiments to be run and fewer animals being used, adhering to the principles of Three Rs (Replacement, Reduction and Refinement) in animal research. The automated homecage system allows many unforeseen issues, such as animal health issues, to be discovered much earlier during the experiments as the recording are transferred all the time. This allows much faster responses to the issues, and potentially saves extra labour and mice if the issues may affect the success of the experiments.

As the hardware and software used in the experiment were both still in prototype stage, the experiment experienced some issues during the recording period. One of the biggest issues being the lighting condition. Each homecage top was equipped with LED lights to provide stable and similar lighting condition across all homecages during the light cycle. However, due to the long recording period, some of the LED lights started to fail during the experiment, resulting in flickering light. When LED lights flickering was observed, the LED lights were disabled through the software application. This means that throughout the experiment, some mice may experience different lighting conditions, which may have short-term or long-term behavioural effects. Since the number of mice affected by the LED issue may not be balanced between the two groups, and the two genotypes may not be affected by the lighting conditions equally, the result of the behavioural analyses can be affected by this. This should be improved as hardware development progresses.

Another difficulty came from network stability. As a real-time long-term online monitoring approach for animal experiments was novel to the university vivarium. The network infrastructure was not optimized for such task. This led to frequent disconnections between the server system and the homecage systems, and that results in frequent experimenter visit and manual fixes. While this was limited to two days a week, it still limited the availability of continuous data. Experimenter visits might also influence the animal behaviours. This can be improved through a dedicated Wi-Fi network. It is also important to ensure software and hardware can maintain their functionalities for long enough before the experimenter visits.

Some further issues came from camera lenses and angles. It is important to keep in mind that the tracking from the video recording is a flat 2D projection. As the homecage systems use wide-angle fisheye lens, the video recordings included a significant amount of distortion. As a result, the distance between two points in the video were warped. In addition, the overhead angle does not provide the depth information, especially at the cage wall and the small elevated platforms on two ends of the cage. These factors both can affect the outcome of the distance-based analyses as the distances calculated in the analyses were not adjusted based on fisheye lens distortion and the missing depth information. There exist novel approaches to create 3D pose estimation from fisheye camera for persons (Zhang, You, Karaoglu, & Gevers, 2022). This can be a future research direction to apply similar technique for both location and pose tracking of mice to mitigate both sources of aberration.

Mouse Tracking

The YOLO-based tracking method has several advantages comparing to other possible tracking techniques. It provides high spatial and temporal resolution. RPi 3B used in this experiment can support recording up to 640x480 resolution and 30 FPS, and more powerful RPi can support even higher. This is much better than Radio Frequency Identification (RFID) based system (Noorshams, 2017), which are also more specialized and expensive. Infrared sensors and piezoelectric system tracks activities with good temporal resolution but offers little to no spatial tracking (Donohue, Medonza, Crane, & O'Hara, 2008). Another video-based tracking method use frame differencing (Singh, Bermudez-Contreras, Nazari, Sutherland, & Mohajerani, 2019). It takes the difference between two frames by subtracting them and tracks the changing area. This approach can be overly sensitive to small changes in the view, such as changes from unstable lighting. A well-trained YOLO model does not have this issue.

While the mouse tracking model trained in this experiment worked well within singly housed mice, it is worth noting that this model would not be able to maintain the identities of multiple mice in a group housing scenario. Since normally lab mice are house in group, tracking multiple mice simultaneously while maintaining their individual identities would be a natural step for future development. One possible solution may be combining YOLO with another object tracking approach called Simple Online and Realtime Tracking (SORT) (Wojke, Bewley, & Paulus, 2017). YOLO and SORT have already been used in object tracking for other objects, such as persons (Bathija & Sharma, 2019). Applying this to mouse tracking in homecage may enable study of social behaviours using automated homecage. As AD mouse models also display changes in

social behaviours as disease progress (Kosel, Munoz, Yang, Wong, & Franklin, 2019), such approach can be beneficial to homecage behavioural study in group housing condition.

The YOLO-based tracking method does require significant time and computing resource for model training and post-processing in dedicated computers with good GPU (Graphics Processing Unit). A much less resource intensive method will be tracking the mice using motion vectors. A motion vector represents the difference in location of the closest matching block (usually in 16x16 pixels) in the previous frame with respect to the block in the current frame. As a result, between every two consecutive frames in a video, there exist a motion vector field representing the changes between the two frames, and it is in a much lower resolution than the original frames. Since motion vectors are already generated and used in the video recording and encoding process by the recording library on the RPi, it can be directly output as a binary file without further processing. In the homecage recording, this motion vector field can be used as an estimation to the amount of mice activity in the homecage, since the changes between two frames are generally result of the movements of the mice. It works similarly to frame differencing, but motion vectors can be computed much faster because of the lower resolution. This allows activity estimation on the RPi in real time even when recording at higher framerates. In addition to the advantage of speed and no need for training or post-processing, motion vector approach can sometimes detect movement undetected by YOLO. For example, when the mice are completely obstructed by the nest, YOLO is not able to locate the mice, so no activity can be detected. Motion vector on the contrary can detect the movement of the nest resulting from the mice movement inside the nest. It can also be used for object

tracking (Kale, Pawar, & Dhulekar, 2015), so it is possible to use it for mouse tracking. However, it would be even harder to maintain identity in a group house setting. It also has the same limitation as frame differencing, which is overly sensitive to small changes like unstable lighting. It was also because of this fact and the LED light issue, even though this approach was implemented in the homecage software system, and the motion vector data was recorded, it was not used in the analysis. Given the relatively low footprint of motion vector recording, combining motion vector and YOLO may provide a more complete picture of the mice behaviours when conditions allow.

Homecage Behavioural Tracking for AD

Human AD patients can display a range of behavioural and neuropsychiatric symptoms (NPSs), such as psychosis, agitation, apathy, depression, and sleep disturbances, each emerging at different time as the disease progress (K. L. Lanctôt, J. Amatniek, S. Ancoli-Israel, S. E. Arnold, C. Ballard, J. Cohen-Mansfield, Z. Ismail et al, 2017). Diagnoses and monitoring of NPSs is one way to predict disease progression for early stages of dementia associated with AD. Similarly, the paradigm of using homecage activity tracking for AD mouse models tries to detect and monitor similar early behavioural changes of the mouse models in the homecage environment. While detecting some of the NPSs such as psychosis, which include hallucinations and delusions, may be challenging in mice, some of them are possible in the homecage setting. Even though the sleep structures are very different between humans and mice, sleep analyses of mice in homecage, such as the detection of more fragmented sleep, are directly associated with the NPS of sleep disturbances. The increase in activity level of the 5xFAD mouse models as they aged could be related to agitation, which includes excessive motor activity.

Detecting and monitoring the emergence and progression of these behavioural changes over time can be useful for behavioural phenotyping and tracking of disease progression. As a result, identifying and profiling these detectable homecage behaviours changes are important.

Homecage recording can be long-term and very rich in detail. While this thesis focuses on analyzing the behaviours through tracking the location of the mice, many more analyses can be done to the data. The 5xFAD is one of the mouse models that can develop seizures as they age. With continuous homecage recording, it is possible to detect and count the number of seizure episodes. The progression of AD also causes the deterioration of the ability to perform complex tasks, which can include behaviours such as nest building. As a result, it is possible to track the progression of the disease by scoring and tracking the score of the nest over time. In addition, employing machine learning techniques can allow classification of mouse behaviours, such as eating, grooming, rearing, etc. With the homecage data and the machine learning approach, it is possible to create patterns of behaviours and investigate the hierarchy of behaviours. While the homecage behaviour paradigm is aiming to minimize human intervention with the mice, some human interactions are unavoidable, such as homecage maintenance. These occasions offer opportunities to observe the behavioural effects after such interactions. Future studies can also extend the recording period to the entire lifetime of the mice, and with different sexes and more strains of mice. The continuous long-term recording can enable investigation of behavioural cycles much longer than a day, such as seasonal cycles and even life cycles (Pernold, Rullman, & Ulfhake, 2021). Performing recording in group housing condition will also allow analyses of social behaviours. As

such, this thesis is not only to verify the homecage behaviour approach for AD, but also to be the beginning point of creating the homecage behavioural profiles of different AD mouse models.

The automated homecage approach can also be highly valuable for use cases such as drug test using AD mouse models. With automated homecage and trackable homecage behavioural metrics, drug tests can not only easily measure the behavioural effect before and after administrative of the drug, but also maintain testing on the same mice for longer periods to study long term effects. This approach allows avoiding the repeated training and testing using behavioural tasks, and the possibility of different tests affecting each other. Adapting the homecage behavioural approach also encourage closer study of how brain pathology of AD mouse models affects homecage behaviours. For example, sleep disturbance and nighttime behaviour changes is one of the earlier NPSs to emerge in human patients (K. L. Lanctôt, J. Amatniek, S. Ancoli-Israel, S. E. Arnold, C. Ballard, J. Cohen-Mansfield, Z. Ismail et al, 2017), but this experiment shows dark cycle behavioural change emerged first in mice. Since mice are nocturnal, intuitively the opposite should be the case as more sleep happens in the light cycle, and the light cycle behaviours should be affected first. The combination of ethological studies using homecage and pathological studies using specialized experiments can provide better insights to the relation between behavioural symptoms and the pathogenesis of AD in the brain. It can also enable AD mouse model developers to create AD mouse models with behavioural traits more closely tied to human AD. In addition, as homecage behavioural tracking is unintrusive and required little direct human intervention, it can be combined with most of other research approaches including specialized behavioural tasks.

Homecage behavioural approach can also easily translate to other species such as other rodents and livestock living in enclosed areas. With proper tracking, it may also be extended to humans.

Small World, Big Data

The homecage behavioural approach takes on the concept of "small world" (Bermudez-Contreras, Sutherland, Mohajerani, & Whishaw, 2022). While homecages are much smaller than and different from the natural habitat of wild mice, they include the entire environment laboratory mice spending the majority or entirety of their lives in and shape the behavioural repertoire of the mice. As a result, while homecage behaviours of lab mice can be different from their counterparts in nature, they are still the natural behaviours of the lab mice. This repertoire of behaviours is affected by diseases such as AD, and with the relative stable and simple environment, they are easier to isolate and study comparing to the behaviours of wild mice in nature.

However, the ability to record continuously for long period of time bring the challenges and opportunities of "big data". The challenges come from the big volumes of data. With 24 cages recording at 10 FPS and 640x480 resolution (12 recording at all time, other 12 recording even other week), the experiment generated close to 10 terabytes (TB) of data. To run experiments with more cages, higher framerates and/or higher resolution, the data throughput will be even higher. This means almost every facet of the experiment requires attention and some level of advanced computer science knowledge from the experimenters. Advanced scheduling and data transfer techniques will be required for the data to be transferred out of the RPi in time. It may also have to involve special

accommodation in the network infrastructures from the network administrator of the institutes, which experimenters may not have direct access to. Dedicated large data storage will be needed to store and process the continuously growing volumes of data, and some compression and archival techniques may be needed. Also, powerful computers and efficient analysis techniques are needed to process as much of the data as quick as possible. Sharing the data can also be challenging because of the large volumes.

However, the large volumes of data provide great opportunities for employing machine learning and deep learning approaches for data analyses. This thesis already utilized object detection system YOLOv8, and some proposed techniques in early section such as SORT and poses/behaviours classification also involves machine learning algorithms. There are still more recently developed machine learning methodologies that can be applied to homecage behaviours analysis. YOLOv8 can perform more than object detection and tracking, but also segmentation, which segment the mice from the rest of the images, classification and pose tracking. DeepLabCut (DLC) is another commonly used deep learning model that can track body parts and postures of mice (Mathis, et al., 2018). Similar to YOLO, it uses supervised learning, which involved manual labeling of frames before training the model. B-SOiD is an unsupervised algorithm that allow classification of behaviours based on postures tracking data, such as those from DLC (Hsu & Yttri, 2021). Since it is unsupervised, no manual labeling is required. Motion Sequencing (MoSeq) is another recently developed approach that can break down animal behaviours into smaller components they called ‘syllables’ (Wiltschko, et al., 2020). Researchers with sufficient knowledge can also develop and train their own machine learning and/or deep learning models for their specific needs. The higher adaptations of

machine learning and deep learning approaches in neuroscience also open new fronts for research, such as using 3D models as synthetic data for training of machine learning model (Bolaños, et al., 2021). This can largely reduce the workload for manual labeling of training data, and also potentially reduce the number of mice needed in experiments. The large variety of machine learning approaches allows investigation in animal behaviours in unprecedented details and with improved reproducibility. Automated homecage approach is to combine the "small world" and "big data" to provide better understand of the rich animal behaviours.

Conclusion

In conclusion, through experiment, this thesis has successfully demonstrated that automated homecage approach is capable of detecting the divergence of behaviours between 5xFAD AD mouse models and non-transgenic mouse models as they aged by analyzing the general activity level, sleep and activity cycles, and excursion behaviours. The observations are consistent with previous research and provide some new insights. Homecage behaviours as a paradigm for AD animal research can provide a different layer of understanding to the disease on top of specialized experiments. Combining automated homecage and machine learning approaches allow behavioural studies of AD at scales and details level not available before.

References

Antunes, M., & Biala, G. (2012). The novel object recognition memory: neurobiology, test procedure, and its modifications. *Cognitive processing, 13*, 93-110.

Bancher, C., Brunner, C., Lassmann, H., Budka, H., Jellinger, K., Wiche, G., . . . Wisniewski, H. M. (1989). Accumulation of abnormally phosphorylated τ precedes the formation of neurofibrillary tangles in Alzheimer's disease. *Brain research, 477*(1-2), 90-99.

Bathija, A., & Sharma, G. (2019). Visual object detection and tracking using yolo and sort. *International Journal of Engineering Research Technology, 8*(11).

Berman, G. J. (2018). Measuring behavior across scales. *BMC biology, 16*, 1-11.

Bermudez-Contreras, E., Sutherland, R. J., Mohajerani, M. H., & Whishaw, I. Q. (2022). Challenges of a small world analysis for the continuous monitoring of behavior in mice. *Neuroscience & Biobehavioral Reviews, 136*, 104621.

Bolaños, L. A., Xiao, D., Ford, N. L., LeDue, J. M., Gupta, P. K., Doebeli, C., . . . Murphy, T. H. (2021). A three-dimensional virtual mouse generates synthetic training data for behavioral analysis. *Nature methods, 18*(4), 378-381.

Brown, L. A., Fisk, A. S., Pothecary, C. A., & Peirson, S. N. (2019). Telling the time with a broken clock: quantifying circadian disruption in animal models. *Biology, 8*(1), 18.

Buslaev, A., Iglovikov, V. I., Khvedchenya, E., Parinov, A., Druzhinin, M., & Kalinin, A. A. (2020). Albumentations: fast and flexible image augmentations. *Information, 11*(2), 125.

Campbell, S. S., & Tobler, I. (1984). Animal sleep: a review of sleep duration across phylogeny. *Neuroscience & Biobehavioral Reviews, 8*(3), 269-300.

Chesler, E. J., Wilson, S. G., Lariviere, W. R., Rodriguez-Zas, S. L., & Mogil, J. S. (2002). Identification and ranking of genetic and laboratory environment factors influencing a behavioral trait, thermal nociception, via computational analysis of a large data archive. *Neuroscience & Biobehavioral Reviews, 26*(8), 907-923.

Crabbe, J. C., Wahlsten, D., & Dudek, B. C. (1999). Genetics of mouse behavior: interactions with laboratory environment. *Science, 284*(5420), 1670-1672.

De Chaumont, F., Ey, E., Torquet, N., Lagache, T., Dallongeville, S., Imbert, A., . . . B. T.-M. (2019). Real-time analysis of the behaviour of groups of mice via a depth-

sensing camera and machine learning. *Nature biomedical engineering, 3*(11), 930-942.

De Felice, F. G., & Munoz, D. P. (2016). Opportunities and challenges in developing relevant animal models for Alzheimer's disease. *Ageing research reviews, 26*, 112-114.

Devi, L., & Ohno, M. (2010). Phospho-eIF2α level is important for determining abilities of BACE1 reduction to rescue cholinergic neurodegeneration and memory defects in 5XFAD mice. *PloS one, 5*(9), e12974.

Donohue, K. D., Medonza, D. C., Crane, E. R., & O'Hara, B. F. (2008). Assessment of a non-invasive high-throughput classifier for behaviours associated with sleep and wake in mice. *Biomedical engineering online, 7*, 1-14.

Drew, V. J., Park, M., & Kim., T. (2023). GABA-Positive Astrogliosis in Sleep-Promoting Areas Associated with Sleep Disturbance in 5XFAD Mice. *International Journal of Molecular Sciences , 24*(11), 9695.

Duhamel, P., & Vetterli, M. (1990). Fast Fourier transforms: a tutorial review and a state of the art. *Signal processing, 19*(4), 259-299.

Duncan, M. J., Smith, J. T., Franklin, K. M., Beckett, T. L., Murphy, M. P., Clair, D. K., . . . O'Hara, B. F. (2012). Effects of aging and genotype on circadian rhythms, sleep, and clock gene expression in APPxPS1 knock-in mice, a model for Alzheimer's disease. *Experimental Neurology, 236*(2), 249-258.

Eilam, D., & Golani, I. (1989). Home base behavior of rats (Rattus norvegicus) exploring a novel environment. *Behavioural brain research, 34*(3), 199-211.

Faraji, J., Ambeskovic, M., Sauter, N., Toly, J., Whitten, K., Lopes, N. A., . . . Metz, G. A. (2022). Sex-specific stress and biobehavioral responses to human experimenters in rats. *Frontiers in neuroscience, 16*, 965500.

Farnworth, B., Innes, J., & Waas, J. R. (2016). Converting predation cues into conservation tools: the effect of light on mouse foraging behaviour. *PloS one, 11*(1), e0145432.

Flanigan, T. J., Xue, Y., Rao, S. K., Dhanushkodi, A., & McDonald, M. P. (2014). Abnormal vibrissa-related behavior and loss of barrel field inhibitory neurons in 5xFAD transgenics. *Genes, Brain and Behavior, 13*(5), 488-500.

Gray, S., & Hurst, J. L. (1995). The effects of cage cleaning on aggression within groups of male laboratory mice . *Animal Behaviour, 49*(3), 821-826.

Hobson, L., Bains, R. S., Greenaway, S., Wells, S., & Nolan, a. P. (2020). Phenotyping in mice using continuous home cage monitoring and ultrasonic vocalization recordings. *Current protocols in mouse biology, 10*(3), e80.

Hsu, A. I., & Yttri, E. A. (2021). B-SOiD, an open-source unsupervised algorithm for identification and fast prediction of behaviors. *Nature communications, 12*(1), 5188.

Iannello, F. (2019). Non-intrusive high throughput automated data collection from the home cage. *Heliyon, 5*(4).

Jankowsky, J. L., & Zheng, H. (2017). Practical considerations for choosing a mouse model of Alzheimer's disease. *Molecular neurodegeneration, 12*, 1-22.

Jocher, G., Chaurasia, A., & Qiu, J. (2023). *YOLO by Ultralytics (Version 8.0.0) [Computer software]*. (Ultralytics) Retrieved 2023, from https://github.com/ultralytics/ultralytics

K. L. Lanctôt, J. Amatniek, S. Ancoli-Israel, S. E. Arnold, C. Ballard, J. Cohen-Mansfield, Z. Ismail et al. (2017). Neuropsychiatric signs and symptoms of Alzheimer's disease: New treatment paradigms. *Alzheimer's & Dementia: Translational Research & Clinical Interventions, 3*(3), 440-449.

Kale, K., Pawar, S., & Dhulekar, P. (2015). Moving object tracking using optical flow and motion vector estimation. *2015 4th international conference on reliability, infocom technologies and optimization (ICRITO)(trends and future directions).* pp. 1-6. IEEE.

Kimura, R., & Ohno, M. (2009). Impairments in remote memory stabilization precede hippocampal synaptic and cognitive failures in 5XFAD Alzheimer mouse model. *Neurobiology of disease, 33*(2), 229-235.

Kleitman, N. (1982). Basic rest-activity cycle—22 years later. *Sleep, 5*(4), 311-317.

Kosel, F., Munoz, P. T., Yang, J. R., Wong, A. A., & Franklin, T. B. (2019). Age-related changes in social behaviours in the 5xFAD mouse model of Alzheimer's disease. *Behavioural Brain Research, 362*, 160-172.

Kundu, P., Torres, E. R., Stagaman, K., Kasschau, K., Okhovat, M., Holden, S., . . . Raber, J. (2021). Integrated analysis of behavioral, epigenetic, and gut microbiome analyses in AppNL-G-F, AppNL-F, and wild type mice. *Scientific Reports, 11*(1), 4678.

Lane, C. A., Hardy, J., & Schott, J. M. (2018). Alzheimer's disease. *European journal of neurology, 25*(1), 59-70.

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature neuroscience , 21*(9), 1281-1289.

Mehla, J., Lacoursiere, S. G., Lapointe, V., McNaughton, B. L., Sutherland, R. J., McDonald, R. J., & Mohajerani, M. H. (2019). Age-dependent behavioral and biochemical characterization of single APP knock-in mouse (APPNL-G-F/NL-G-F) model of Alzheimer's disease. *Neurobiology of Aging, 75*, 25-37.

Morris, R. G. (1981). Spatial localization does not require the presence of local cues. *Learning and motivation, 12*(2), 239-260.

Noorshams, O. (2017). Automated systems to assess weights and activity in group housed mice. *Doctoral dissertation, University of British Columbia*.

Oakley, H., Cole, S., Logan, S., Maus, E., Shao, P., Craft, J., . . . Berry, R. (2006). Intraneuronal beta-amyloid aggregates, neurodegeneration, and neuron loss in transgenic mice with five familial Alzheimer's disease mutations: potential factors in amyloid plaque formation. *Journal of Neuroscience, 26*(40), 10129-10140.

Pernold, K., Rullman, E., & Ulfhake, B. (2021). Major oscillations in spontaneous home-cage activity in C57BL/6 mice housed under constant conditions. *Scientific Reports, 11*(1), 4961.

Piaceri, I., Nacmias, B., & Sorbi, S. (2013). Genetics of familial and sporadic Alzheimer's disease. *Frontiers in Bioscience-Elite, 5*(1), 167-177.

Puzzo, D., Lee, L., Palmeri, A., Calabrese, G., & Arancio, O. (2014). Behavioral assays with mouse models of Alzheimer's disease: practical considerations and guidelines. *Biochemical pharmacology, 88*(4), 450-467.

Ramanathan, C., Stowie, A., Smale, L., & Nunez, A. A. (2010). Phase preference for the display of activity is associated with the phase of extra-suprachiasmatic nucleus oscillators within and between species. *Neuroscience , 170*(3), 758-772.

Richardson, C. A. (2015). The power of automated behavioural homecage technologies in characterizing disease progression in laboratory mice: A review. *Applied Animal Behaviour Science, 163*, 19-27.

Richter, S. H. (2020). Automated home-cage testing as a tool to improve reproducibility of behavioral research? *Frontiers in Neuroscience, 14*, 383.

Sethi, M., Joshi, S. S., Webb, R. L., Beckett, T. L., Donohue, K. D., Murphy, M. P., . . . Duncan, M. J. (2015). Increased fragmentation of sleep–wake cycles in the 5XFAD mouse model of Alzheimer's disease. *Neuroscience, 290*, 80-89.

Sherwin, C. (1998). Voluntary wheel running: a review and novel interpretation. *Animal Behaviour, 56*(1), 11-27.

Singh, S., Bermudez-Contreras, E., Nazari, M., Sutherland, R. J., & Mohajerani, M. H. (2019). Low-cost solution for rodent home-cage behaviour monitoring. *Plos one, 14*(8), e0220751.

Sturman, O., Germain, P.-L., & Bohacek, J. (2018). Exploratory rearing: a context-and stress-sensitive behavior recorded in the open-field test. *Stress, 21*(5), 443-452.

Van Dam, D., & Deyn, P. P. (2011). Animal models in the drug discovery pipeline for Alzheimer's disease. *British journal of pharmacology, 164*(4), 1285-1300.

van der Linde, R. M., Dening, T., Matthews, F. E., & Brayne, C. (2014). Grouping of behavioural and psychological symptoms of dementia. *International journal of geriatric psychiatry, 29*(6), 562-568.

Voelkl, B., & Würbel, H. (2016). Reproducibility crisis: are we ignoring reaction norms? *Trends in Pharmacological Sciences, 37*(7), 509-510.

Wiltschko, A. B., Tsukahara, T., Zeine, A., Anyoha, R., Gillis, W. F., Markowitz, J. E., . . . Datta, S. R. (2020). Revealing the structure of pharmacobehavioral space through motion sequencing. *Nature neuroscience, 23*(11), 1433-1443.

Wojke, N., Bewley, A., & Paulus, D. (2017). Simple online and realtime tracking with a deep association metric. *2017 IEEE international conference on image processing (ICIP).* pp.3645-3649. IEEE.

World Health Organization. (2021). *Global status report on the public health response to dementia.*

Xu, W., Tan, L., Wang, H.-F., Jiang, T., Tan, M.-S., Tan, L., . . . Yu, J.-T. (2015). Meta-analysis of modifiable risk factors for Alzheimer's disease. *Journal of Neurology, Neurosurgery & Psychiatry, 86*(12), 1299-1306.

Zhang, Y., You, S., Karaoglu, S., & Gevers, T. (2022). Multi-person 3D pose estimation from a single image captured by a fisheye camera. *Computer Vision and Image Understanding, 222*, 103505.