INTRODUCTION

Preclinical to Clinical Translation Gap in Alzheimer's Therapy Development

Positive findings of therapeutics and interventions from studies conducted in Alzheimer's disease (AD) animal models are often not translated to effective treatments. As comprehensively reviewed by Zahs and Ashe, 2010, more than 200 interventions have been reported to be efficacious in ameliorating pathology and cognitive deficits in transgenic AD animal models, yet none have proven effective in human trials. Detailed analyses of AD clinical trials of candidate drugs considered safe and effective, based on successful preclinical animal studies conducted between 1998 and 2014, were performed by Cummings et al., 2014 and PhRMA, 2015 and reported a very high attrition rate of more than 99%. During this period 244 candidate therapeutics were tested in 413 clinical trials (phases I-III) and only one received FDA approval as a disease-modifying therapeutic – memantine. Systematic reviews and meta-analyses of preclinical animal studies in AD highlight the lack of methodological rigor and inadequate reporting practices as contributive to the preclinical to clinical translation gap in AD therapy development (Shineman et al., 2011, Landis et al., 2012, Snyder et al., 2016, Percie du Sert et al., 2020).

To combat the challenges presented by the poor translation of AD animal research to effective treatments, the National Institute on Aging (NIA) has convened a number of AD research summits (2012, 2015, 2018, 2021) aimed at identifying and addressing critical knowledge gaps and proposing ways to utilize emerging technologies to accelerate treatments and preventative strategies for people at all stages of the disease. These summits have brought together leading experts on AD from academia, industry, non-profit organizations, and advocacy groups to develop research priorities and strategies needed to accelerate the development of successful therapies for AD. Also identified during these summits were infrastructure and partnerships necessary to increase the likelihood that preclinical therapeutic development efforts for AD will translate to success in the clinic.

The AD research summits delivered recommendations that served as the foundation for developing the AD/ADRD Research Implementation Milestones. These milestones encompass the entire AD research landscape including basic, translational, clinical and health services research and detail specific measures and success criteria towards the development of effective treatments for AD. One such milestone (4.B in particular) tasked the NIA to create infrastructure and resources for the development of standardized and rigorous methods for preclinical efficacy testing, including web-based resources for transparent reporting of both positive and negative findings. Specifically, the milestone recommended the development of a publicly available database that housed experimental design and analyses relating to preclinical testing of candidate therapeutic agents in AD animal models and identified critical experimental design elements and methodology that impact rigor and reproducibility of published studies.

AlzPED: Mission, Scope, and Capabilities

Launched in 2016 by the NIA and the NIH Library, the Alzheimer's Disease Preclinical Efficacy Database (AlzPED) incorporates all of these recommendations and is a key component of NIA's new translational infrastructure for AD/ADRD research. AlzPED is a searchable, publicly available knowledgebase hosting curated summaries from nearly 1300 published animal model studies testing candidate therapeutics in AD, aiming to illuminate key elements of experimental design and reporting practices of preclinical efficacy testing studies for researchers, funding agencies, and the public. These curated summaries include information about AD animal models, therapeutic agents, therapeutic targets, outcome measures, related clinical trials, patents, and study design.

The database is missioned to improve transparency in reporting, increase awareness of the need for greater rigor in study design, and identify critical experimental design elements and methodology missing from studies that make them susceptible to over-interpretation and reduce their reproducibility and translational value. Through these capabilities, AlzPED is intended to guide the development and implementation of strategies and recommendations for standardized best practices to achieve rigorous preclinical testing of AD candidate therapeutics.

METHODS

Article Selection

AlzPED provides platforms for the submission of both published as well as unpublished studies testing candidate therapeutics in AD animal models, particularly unpublished studies that describe negative results. Published studies are collected from databases like PubMed and Embase using key word search strings specific to preclinical therapeutic testing in AD animal models. Unpublished studies, particularly those describing negative results, will be obtained directly from researchers testing candidate therapeutic agents in AD animal models. Efforts to procure these data are continuous and ongoing.

Article Curation

Prior to publication in the database, each study is meticulously curated by 2 experts in AD research for bibliographic details, funding source, study goals and principal findings, data on relevant translational criteria like therapy type, therapeutic agent, therapeutic target, animal models, and AD-related outcome measures. Studies are also assessed for scientific rigor using a "Rigor Report Card" (Figure 1) which identifies the experimental design elements which were reported as well as those missing from the study. Each curated study provides additional relevant information about therapeutic targets, therapeutic agents, and animal models through freely accessible external databases. These databases include Open Targets, Pharos, and Agora, which provide comprehensive and integrated insight into potential therapeutic targets for AD; databases like PubChem and DrugBank which describe molecular, chemical, structural and other identifiers for the therapeutic agents being tested; information on clinical trials and patents associated with the therapeutic agent is gleaned from ClinicalTrials.gov and Google Patents; and finally, Alzforum provides detailed descriptions of the animal models used.

Data Analysis and Statistics

All data in the database are publicly available and downloadable from the AlzPED "Search Results" page. Data were analyzed across the translational research criteria of therapy type, therapeutic agent, animal models, and outcome measures. Data are presented as percentages derived from the totality of studies curated in AlzPED. Scientific rigor in study design and methodology was evaluated using the Rigor Report Card consisting of a standardized set of 24 experimental design elements (Figure 1) and results are presented as percentage of studies reporting each of the 24 experimental design elements (Figure 2). Of these 24, 9 core study design elements have been identified as critical for ensuring rigor and reproducibility (Figure 3). Reporting trends in the 9 critical core experimental design elements were evaluated over 5-year spans from 2000 to 2022. Results are presented as "number of studies that reported" vs "number of studies that did not report" the core experimental design elements and analyzed using Chi square tests (Figure 4). Comparisons of reporting trends in the 9 core experimental design elements between NIH-funded and non-NIH funded studies are assessed using Pearson Chi square and Fisher exact tests and expressed as percentages (Figures 6 & 7).

Reporting trends in the experimental design elements were also evaluated based on the impact factor of the journal in which the curated preclinical study is published as well as relative number of citations per year for the study. Current journal impact factors were obtained from Journal Citation Reports. Curated studies were categorized into 4 groups based on current journal impact factor values. Group 1 includes studies published in journals with impact factors below 3, and those published in journals with impact factors between 3 and 4.99, or between 5 and 9.99 are sorted in Groups 2 and 3 respectively. Studies published in high impact journals with impact factors greater than 10 are included in Group 4 (Figure 8). The number of citations for each curated study was obtained from iCite. iCite is a tool to access a dashboard of bibliometrics for published scientific papers developed by the Office of Portfolio Analysis at the NIH. Relative number of citations for each curated study was calculated by dividing the total number of citations for that study by the number of years since publication. For example, for a study published in 2020, the total number of citations for that study was divided by 2, or for a study published in 2019, the total number of citations for that study was divided by 3, and so on. These studies were categorized into 3 groups based on the relative number of citations per year. Curated studies with less than 3 relative number of citations per year are sorted into Group 1, those with relative number of citations per year between 3 and 7 or those with relative number of citations per year greater than 7 are sorted into Groups 2 and 3 respectively. Data are expressed as Mean ± SEM and assessed for significance using unpaired t tests (Figure 9).

RESULTS

This growing knowledge platform currently houses 1402 preclinical efficacy studies published between 1996 and 2022 (Table 1). Efforts are underway to expand the database further and balance the number of studies curated based on the year of publication.

YEAR	# OF PAPERS CURATED	YEAR	# OF PAPERS CURATED	YEAR	# OF PAPERS CURATED
1996	1	2009	65	2020	84
1999	1	2010	72	2021	73
2000	15	2011	64	2022	56
2001	19	2012	64		
2002	27	2013	73		
2003	33	2014	73		
2004	58	2015	76		
2005	53	2016	71		
2006	61	2017	85		
2007	60	2018	77		
2008	60	2019	81		

Table 1: The table shows the number of preclinical efficacy studies collected, curated and available in the database and the year of publication. The database currently houses curated summaries from 1402 published studies.

Experimental Design Elements

AlzPED's Rigor Report Card (Figure 1) provides a standardized set of 24 study design elements that are specifically tailored for rigorous preclinical testing of candidate therapeutics in AD animal models. The Rigor Report Card further defines a succinct set of 9 core design elements that are critical for scientific rigor and reproducibility, without which the reliability of results cannot be assessed (Figure 3). These include design elements that report on power calculation and sample size, blinding for treatment allocation as well as analysis of outcome measures, random allocation of intervention, age, sex, and genetic background of the animal model used, and whether the study has been balanced for sex. Also included are eligibility criteria like inclusion and exclusion criteria that define which animals are eligible to be enrolled in the study, premature deaths and drop-outs, and inclusion of author conflict of interest statement. Further, design elements that characterize the therapeutic agent being tested such as pharmacokinetic, pharmacodynamic, ADME and toxicology measures, dose, and formulation of the therapeutic agent as well as well as treatment paradigms such as route, frequency and duration of administration, and plan for statistical analysis of study results are also included. These are all common elements of clinical trial study design (van der Worp et al., 2010, Shineman et al., 2011, Landis et al., 2012, Snyder et al., 2016, Percie du Sert et al., 2020). The ARRIVE guidelines provide clear definitions and examples for each study design element.

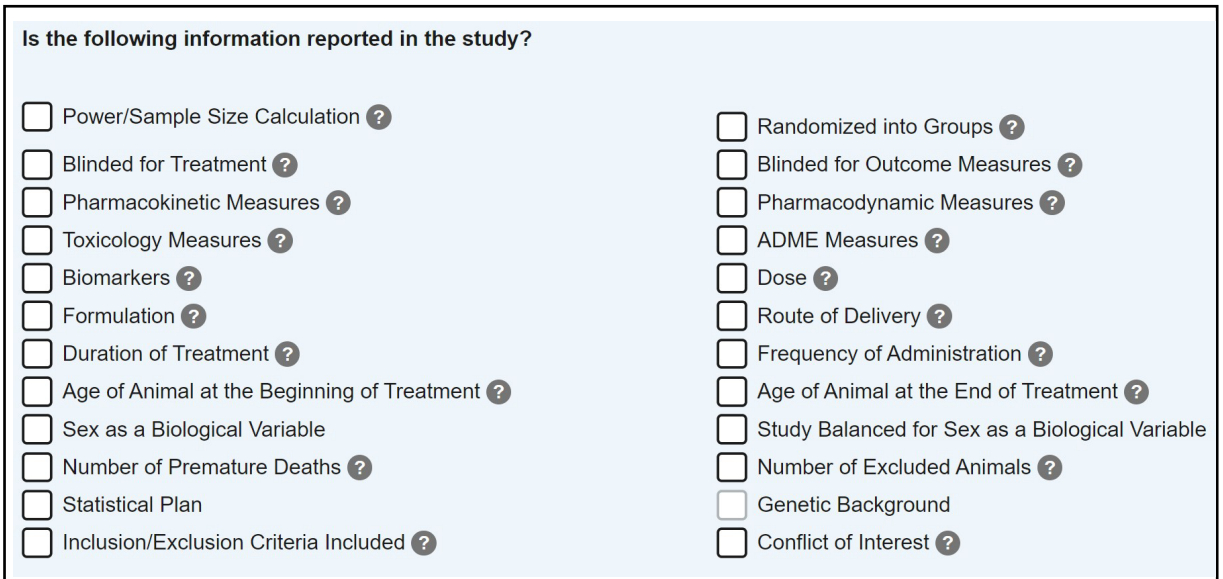

Figure 1: AlzPED is designed to monitor the scientific rigor of curated studies with a "Rigor Report Card" consisting of a standardized set of 24 experimental design elements recommended by expert advisory groups for preclinical therapeutic studies in animals.

Scientific rigor of each curated study is monitored using AlzPED's Rigor Report Card, which identifies experimental design elements reported in a study as well as those missing from the study. Evaluation of these rigor report cards demonstrate considerable disparities in reporting the 24 experimental design elements (Figure 2). Design elements like dose and formulation of the therapeutic agent being examined, treatment paradigms like route, frequency and duration of administration, age of the animals used in the study, and plan for statistical analysis of study results are consistently reported by at least 95% of the 1402 curated studies. However, critical elements of methodology like power calculation, blinding for treatment allocation, as well as analysis of outcome measures, random allocation of

intervention, inclusion and exclusion criteria, and whether the study has been balanced for sex are significantly under-reported, being reported by fewer than 35% of the 1402 curated studies (Figure 3).

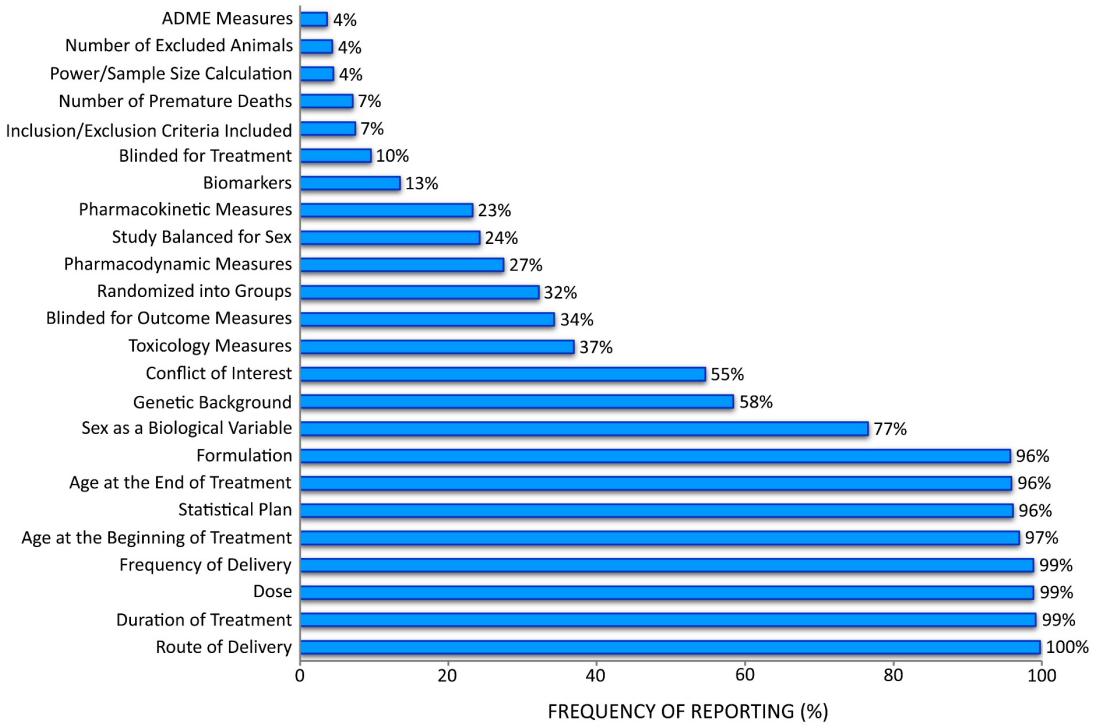

Figure 2: Percentage of studies reporting the 24 experimental design elements identified to improve rigor and reproducibility, derived from 1402 published preclinical studies curated in AlzPED.

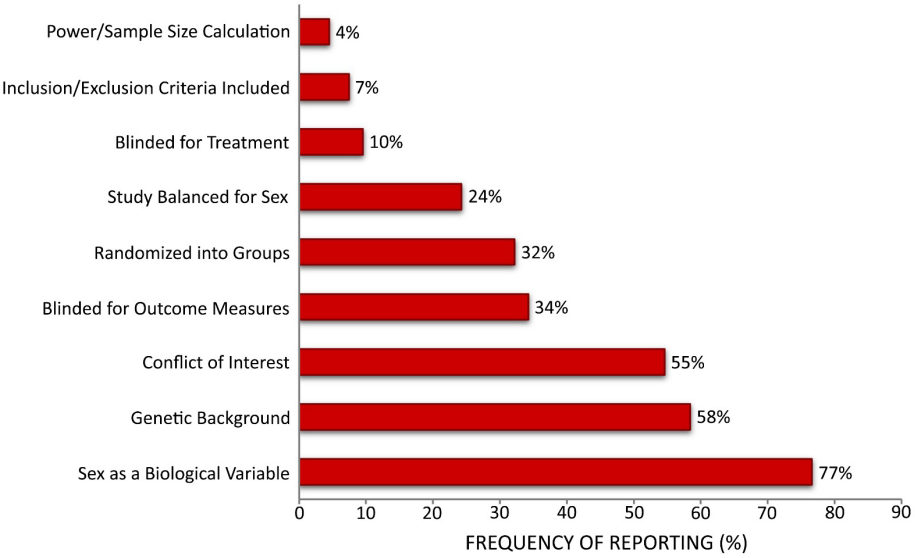

Figure 3: Percentage of studies that reported 9 core experimental design elements derived from 1402 published preclinical studies curated in AlzPED.

Reporting trends for the 9 critical core experimental design elements were evaluated over 5-year spans from 2000 to 2022 (Figure 4). Our analysis demonstrates an incremental improvement over consecutive 5-year spans in the proportion of studies reporting several critical elements like blinding for outcome analysis, randomized allocation of therapeutic agent, genetic background as well as sex of the animal model used in the study, and author conflict of interest statement. However, there is little improvement in the proportion of studies reporting power calculation, blinding for treatment allocation, inclusion and exclusion criteria, and whether the study has been balanced for sex. Of note, NIH-issued 2 rigor-related policies – (i) NOT-OD-11-109 issued in 2011 requires transparency in reporting financial conflicts of

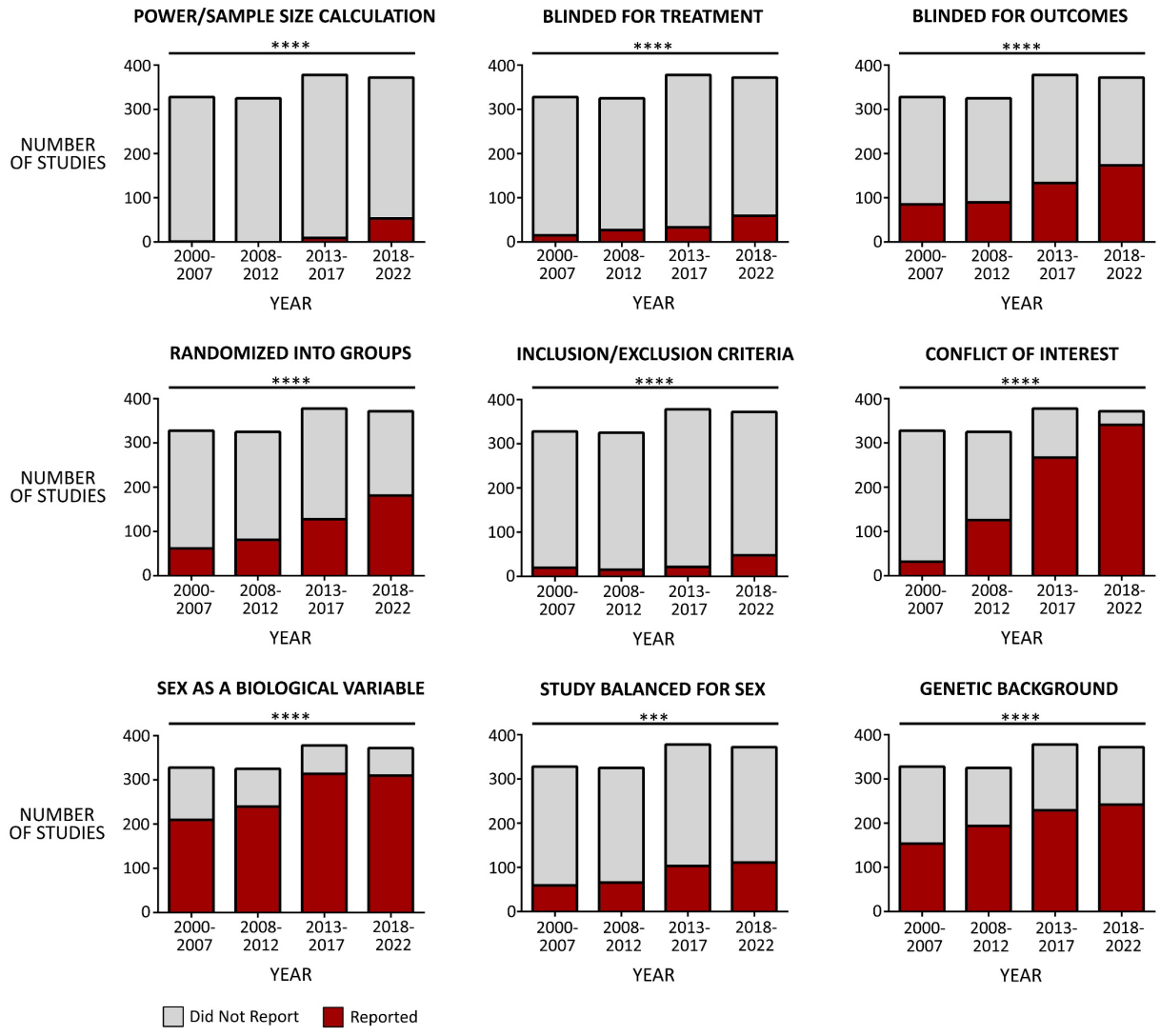

Figure 4: Graphs show reporting trends for the 9 critical core experimental design elements evaluated over 5-year spans from 2000 to 2022. Data analyzed using Chi square test; ***p<0.001, ****p<0.0001. Data presented as number that reported Vs number that did not report core experimental design elements, calculated from 328, 325, 378 and 372 curated studies published between 2000-2007, 2008-2012, 2013-2017 and 2018-2022 respectively.

interest, and (ii) NOT-OD-15-102 was issued in 2015 to emphasize the consideration of sex as a biological variable. Enforcement of these policies clearly improved the reporting of these core experimental design elements.

Further evaluation of the reporting trends in the 9 critical core experimental design elements demonstrates that few studies report more than 5 core design elements, most studies reporting only 2-4 core design elements (Figure 5, top). However, studies published between 2018 and 2022 report between 3-5 core design elements compared with those published between 2000 and 2007 which report 1-2 core design elements (Figure 5, bottom).

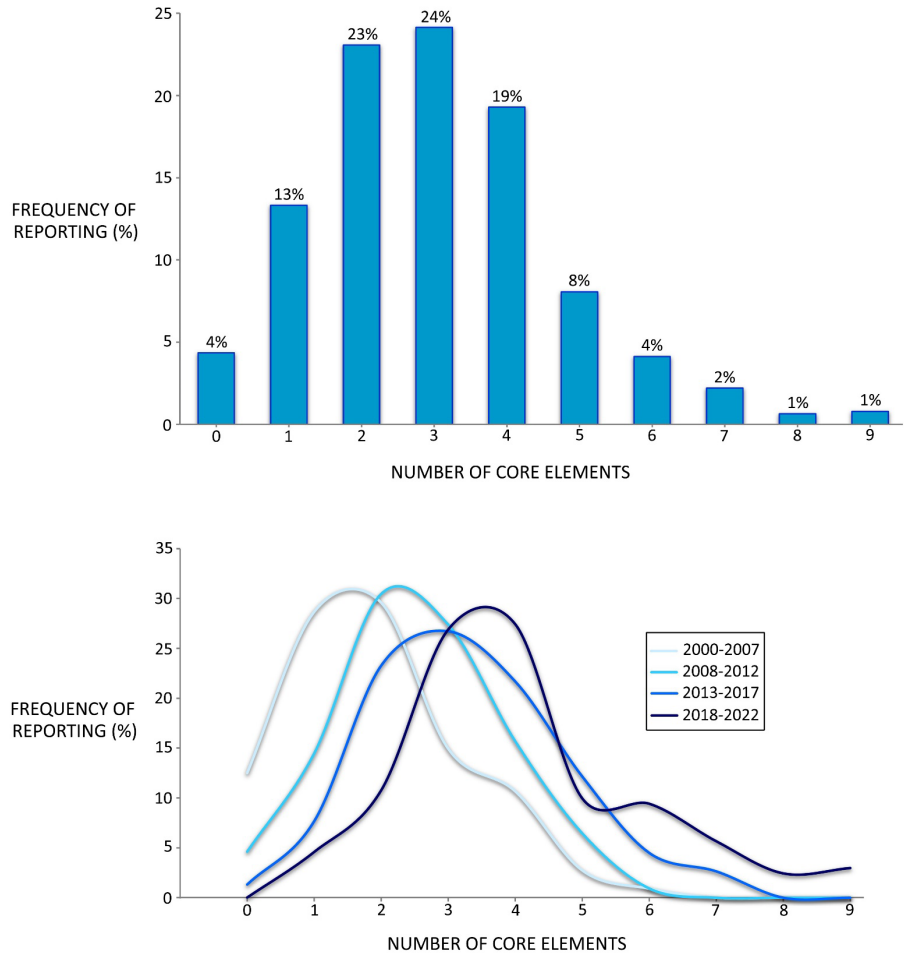

Figure 5: Top: Percentage of studies published between 2000-2022 that report 0-9 core experimental design elements, derived from 1402 published preclinical studies curated to AlzPED. **Bottom:** Graphs show reporting trends of core elements during 5-year spans, derived from 328, 325, 378 and 372 curated studies published between 2000-2007, 2008-2012, 2013-2017 and 2018-2022 respectively.

NIH-issued policies to enhance the rigor, reproducibility and translatability of research place significant emphasis on rigorous scientific method, study design, and consideration of biological variables such as sex. Analysis of study design and methodology of NIH-funded published studies curated in AlzPED demonstrates a positive impact of these policies. NIH-funded studies show significantly improved reporting of power/sample size calculation, blinding for outcome measures, inclusion/exclusion criteria and balancing the study for sex (Figure 6). Non-NIH funded studies include those funded by US federal and state organizations, nonprofit organizations, pharmaceutical companies, the European Union, non-EU European governments, governments of Canada, Mexico, Brazil, China, Japan, South Korea, India, Iran, Australia and others, and University start-ups.

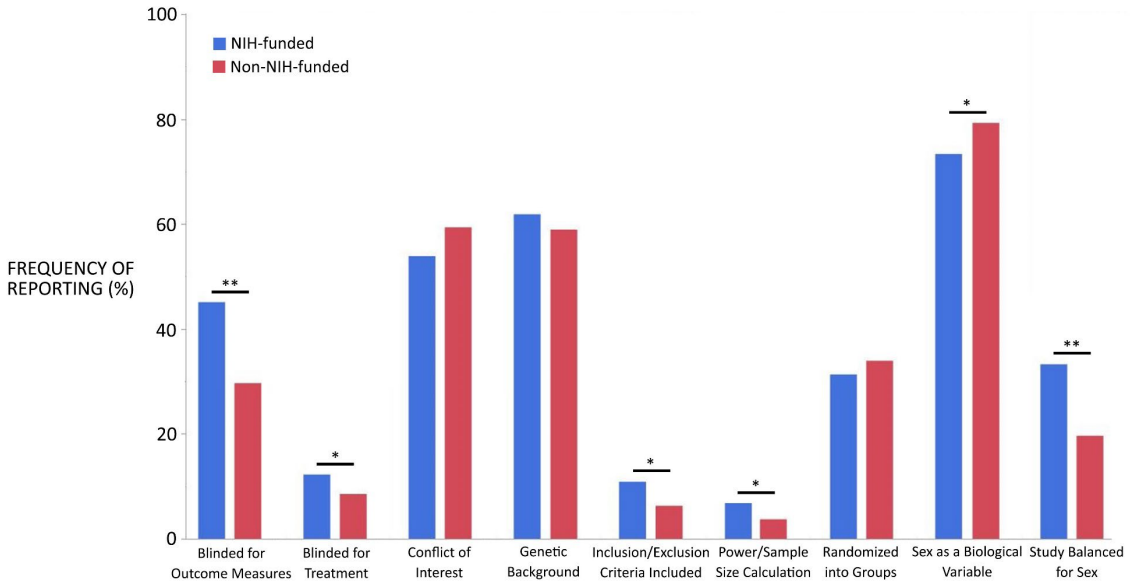

Figure 6: Comparison of reporting trends in 9 core experimental design elements between NIH-funded and non-NIH funded studies. Data are presented as percentages calculated from 514 NIH-funded studies and 748 studies funded by non-NIH sources. Data are analyzed using the 2-tailed Fisher Exact test, *p<0.05 and **p<0.01.

Studies funded exclusively by the NIH show significantly improved reporting of inclusion/exclusion criteria and balancing the study for sex (Figure 7) compared with those supported by specific non-NIH funding sources that include the European Union, nonprofit organizations, pharmaceutical companies, governments of China and other south and east Asian countries (Japan, South Korea, India, Hong Kong and Taiwan).

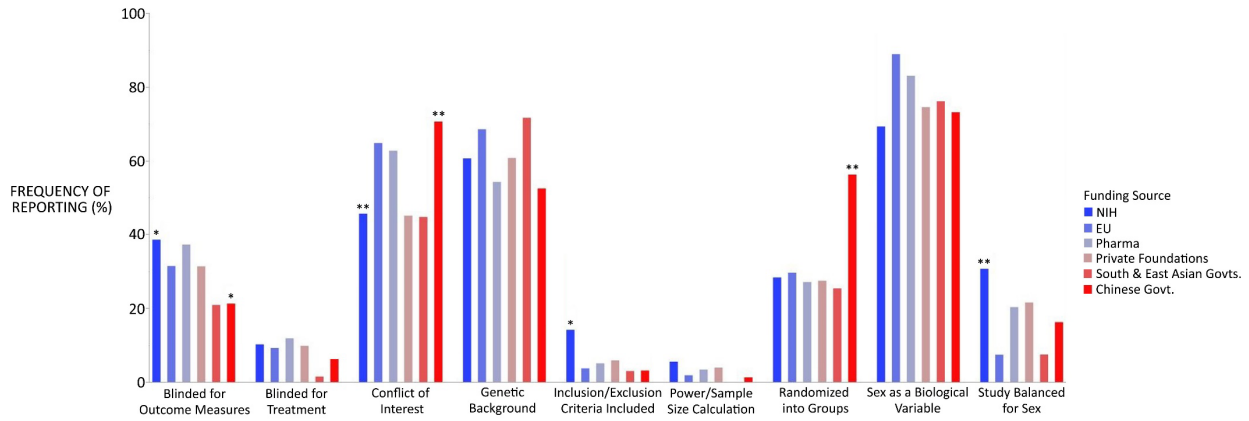

Figure 7: Comparison of reporting trends in 9 core experimental design elements between NIH-funded and non-NIH funded studies. Data are presented as percentages calculated from 127 (exclusively) NIH-funded studies, 54 studies funded by the European Union, 59 pharmaceutical industry-funded studies, 51 studies funded by private foundations, 67 studies funded by south and east Asian governments and 160 studies funded by the Chinese government. Studies included in the south and east Asian government funding category were funded by Japanese, South Korean, Indian, Taiwanese and Hong Kong governments. Data are analyzed using the Pearson Chi square test, *p<0.05 and **p<0.001 presented as percentage reported.

The lack of rigor and reproducibility of research findings in the scientific publishing arena is well documented. A systematic review of the rigor and translatability of highly cited animal studies published in leading scientific journals (including Science, Nature, Cell, and others) demonstrated a lack of scientific rigor in study design. A comprehensive review of reporting trends for critical experimental design elements in highly cited studies published in leading journals that are curated in AlzPED revealed a similar pattern of poor study design and reporting practices. These analyses are described in greater detail in the next segment.

Reporting trends for the 9 core and all 24 experimental design elements were evaluated based on the impact factor of the journal in which the curated preclinical study was published (Figure 8). These studies were categorized into 4 groups based on 2021 journal impact factor values. Curated studies published in journals with impact factors below 3 were sorted in Group 1, and those published in journals with impact factors between 3 and 4.99, or between 5 and 9.99 were sorted in Groups 2 and 3 respectively. Studies published in high impact journals with impact factors greater than 10 were sorted in Group 4. While t-tests show that there are statistically significant differences in reporting the 9 core elements as well as all 24 elements of experimental design between these four groups, overall, the data demonstrate poor reporting practices irrespective of journal impact factor.

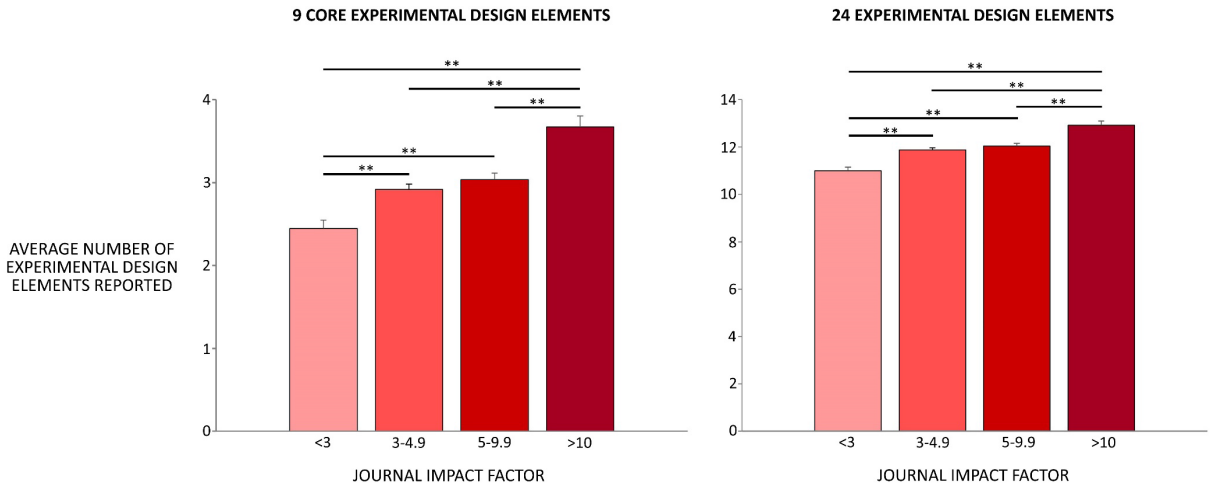

Figure 8: Reporting trends for the 9 core and all 24 experimental design elements based on journal impact factor. Data are presented as Mean ± SEM and analyzed using two-tailed t-tests, **p<0.001 from 159, 580, 462 and 203 curated studies published in journals with impact factors below 3, between 3 and 4.9, between 5 and 9.9, and above 10, respectively.

Reporting trends for the 9 core and all 24 experimental design elements were evaluated based on the relative number of citations per year of each curated study published between 1996 and 2022 (Figure 9). Relative number of citations for each curated study was calculated by dividing the total number of citations for that study by the number of years since publication. For example, for a study published in 2020, the total number of citations for that study was divided by 2, or for a study published in 2019, the total number of citations for that study was divided by 3, and so on. These studies were categorized into 3 groups based on the relative number of citations per year. Curated studies with less than 3 relative number of citations per year were sorted into Group 1, those with relative number of citations per year between 3 and 7 or those with relative number of citations per year greater than 7 were sorted into Groups 2 and 3 respectively. While t-tests show that there are statistically significant differences in reporting the 9 core elements as well as all 24 elements of experimental design between these three groups, overall, the data demonstrate poor reporting practices irrespective of relative number of citations per year.

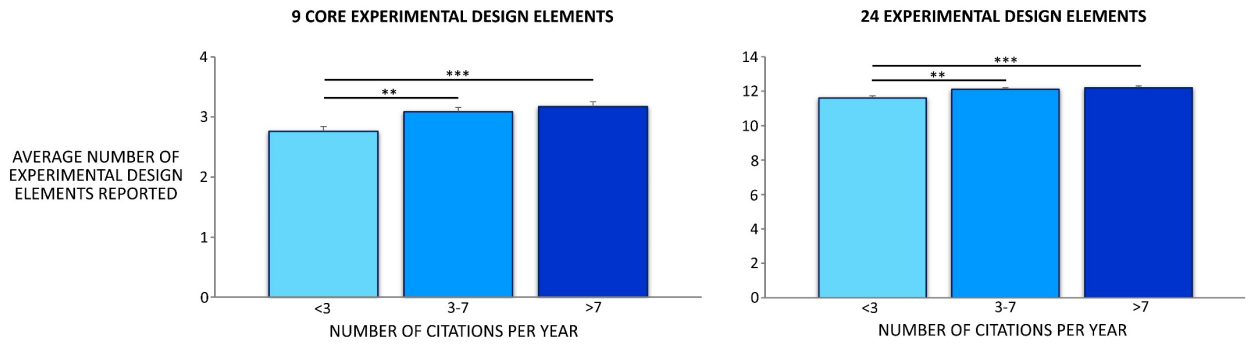

Figure 9: Reporting trends for the 9 core and all 24 experimental design elements based on number of citations per year. Data are presented as Mean ± SEM and analyzed using two-tailed t-tests, **p<0.001, and ***p<0.0001 from 433, 507 and 462 curated studies with number of citations per year below 3, between 3 and 7, and above 7, respectively.

In addition, the reporting trends for the 9 core and all 24 experimental design elements were also analyzed in terms of their clinical impact (Figure 10). Clinical impact reflects the influence a publication has in informing clinical studies. An article is considered to have clinical impact if it was cited by any one of the following: a Clinical Trial that has results (from clinicaltrials.gov), Clinical Study, Clinical Trials (Phases I-IV), Adaptive Clinical Trial, Controlled Clinical Trial, Randomized Controlled Trial, Clinical Trial Protocol, Observational Study, Guideline, or Practice Guideline. For successful translation of candidate therapeutics to the clinic, preclinical studies informing clinical investigations should be conducted with highest rigor. However, in analyzing the reporting trends of the core 9 experimental design elements indicative of rigor using t-tests, studies that have clinical impact were found to report fewer core experimental design elements on average than studies that do not.

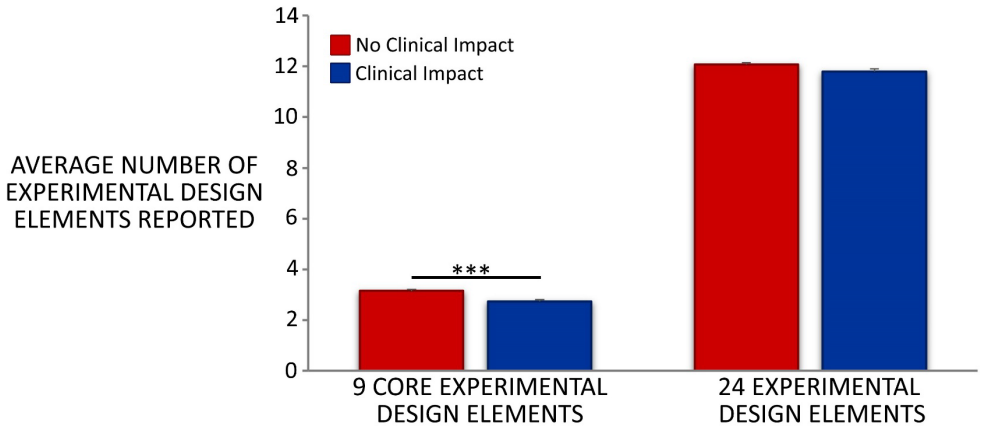

Figure 10: Reporting trends for the 9 core and all 24 experimental design elements comparing articles with and without clinical impact. Data are presented as Mean ± SEM and analyzed using two-tailed t-tests, ***p<0.0001 from 451 studies with clinical impact and 951 studies without clinical impact.

Cumulatively, these data demonstrate the serious deficiencies in reporting critical elements of study design and methodology in the 1400 curated studies analyzed, even those published in high impact factor journals as well those that are highly cited. These deficiencies diminish the scientific rigor, reproducibility, and predictive value of preclinical therapeutic studies done in AD animal models. In light of these results, adoption and implementation of a standardized set of best practices and study design guidelines will very likely improve the predictive validity of preclinical therapeutic studies done in AD animal models. This measure will likely promote the effective translation of preclinical candidate therapeutic testing to the clinic. To this end, several meetings and workshops have been held between the NIH and journal publishers to discuss issues of increasing reproducibility, improving rigor of research findings, identifying common opportunities to enhance rigor in study design, and supporting research that is reproducible and transparent. For example, notice NOT-OD-11-09 issued by the NIH in 2011 requires transparency in reporting financial conflicts of interest. Similarly, most publishers now require investigators to report financial conflicts of interest. Enforcement of these requirements clearly improved reporting this design element. It is therefore imperative that the NIH, other federal funding agencies, private foundations and scientific journal publishers require investigators to follow these accepted best practices and study design guidelines to ensure that funded as well as published research are sufficiently rigorous, transparent, and reproducible.

Therapeutics

A diverse array of therapeutic agents and targets are reported in the 1298 studies curated in AlzPED. The database catalogues 1206 novel therapeutic agents into 16 distinct categories (Figure 11) based on agent source (natural product or synthetic), molecular structure (biologic or small molecule), chemical nature (peptide, nucleic acid, or hormone) and mechanism of action (immunotherapy – active or passive).

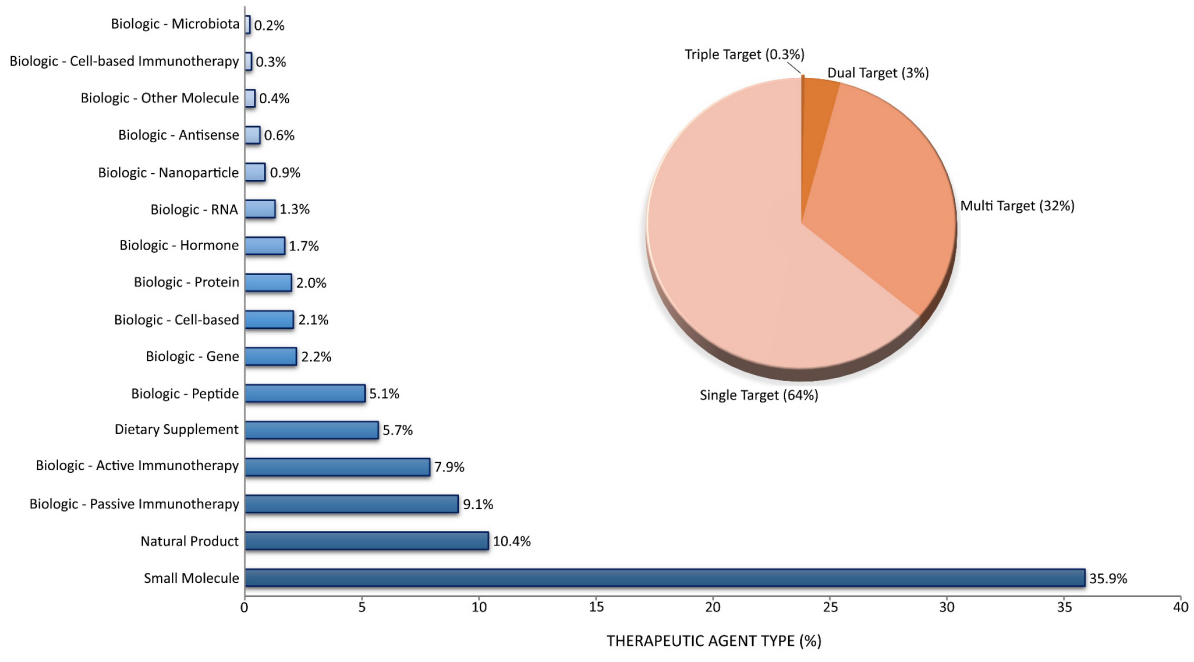

Figure 11: The graph shows the diverse types of therapeutic agents tested, derived from 1402 published preclinical studies curated to AlzPED. **Inset:** Pie chart shows percentage of single, dual, triple and multi target therapeutic agents tested.

Notably, numerous therapeutic agents demonstrate varying extents of anti-inflammatory, antioxidant, beta amyloid-reducing, neuroprotective and cognition enhancing properties and are categorized as multi target therapeutics. Agents that have multiple documented targets are also categorized as multi target therapeutics. The database also describes several therapeutic agents that have dual or triple target specificities (Figure 11, inset). Examples of dual/triple specificity therapeutic agents include dual or triple receptor agonists or antagonists, metal chelators and ionophores.

Currently, AlzPED stores information on 274 therapeutic targets that aim to reduce beta amyloid and tau-related pathology and address disease-associated inflammation, oxidative stress, metabolic, synaptic, and behavioral dysfunction. These assorted targets are categorized into amyloidogenic proteins, tau protein, non-amyloid proteins, enzymes, channels, receptors and transporters, metal ions, and free radicals (Figure 12).

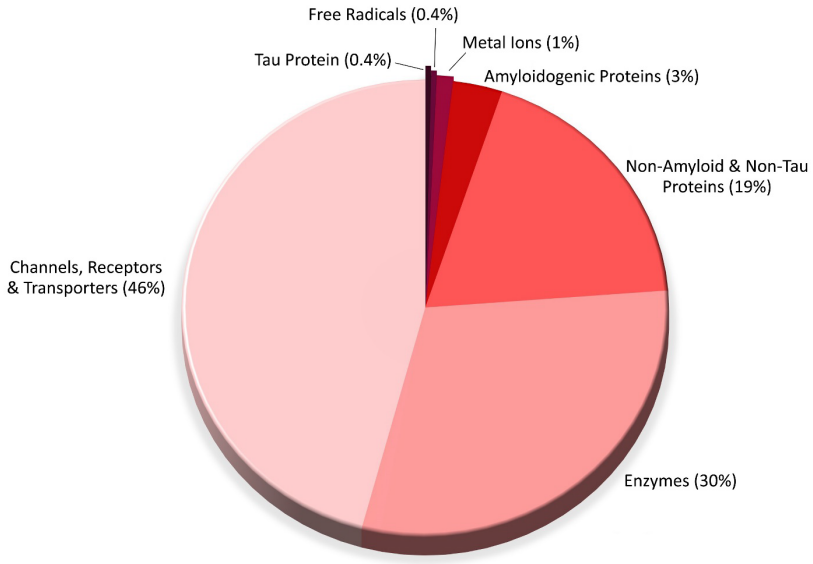

Figure 12: The graph shows the diverse categories of therapeutic targets, derived from 1402 published preclinical studies curated to AlzPED.

AlzPED also includes 28 therapeutic targets (Table 2) from a list of more than 500 nascent drug targets that have been nominated by researchers from the NIA's Accelerating Medicines Partnership in Alzheimer's Disease (AMP-AD) Consortium.

Therapeutic Target	Gene
alpha Synuclein	SNCA
Angiotensin I Converting Enzyme (ACE)	ACE
Apolipoprotein E (ApoE)/Apolipoprotein E4 (ApoE4)	APOE
Brain-Derived Neurotrophic Factor (BDNF)	BDNF
CD33	CD33/SIGLEC3
Colony Stimulating Factor 1 Receptor (CSF1R)	CSF1R/CSFR
Corticotropin-Releasing Factor Receptor 1 (CRFR1)	CRHR1/CRFR/CRF1/CRFR1
Excitatory Amino Acid Transporter 2 (EAAT2)	SLC1A2/EAAT2
Extracellular Signal-Regulated Kinase 2 (ERK2)	ERK2/MAPK1
Glucocorticoid Receptors	NR3C1
Glycoprotein 130/Interleukin 6 Signal Transducer (gp130/IL6ST)	GP130/IL6ST
Histone Deacetylase 1 (HDAC1)	HDAC1
Insulin Receptor	INSR
Insulin-Like Growth Factor 1 Receptor (IGF1R)	IGF1R
Myelin Oligodendrocyte Glycoprotein (MOG)	MOG
Myelin Proteolipid Protein (PLP)	PLP1
Neurotrophin Receptor p75 (p75NTR)	NGFR
Peroxisome Proliferator-Activated Receptor alpha (PPAR alpha)	PPARA
Programmed Cell Death Protein 1 (PD-1)	PDCD1/PD1
Prosaposin	PSAP
Retinoid X Receptor alpha (RXR alpha)	RXRA
Signal Transducer and Activator of Transcription 3 (STAT3)	STAT3
Specificity Protein 1 (Sp1)	SP1
Sphingosine-1-Phosphate Receptor	S1PR1
TAR DNA-Binding Protein 43 (TDP-43)	TARDBP
Triggering Receptor Expressed on Myeloid Cell 2 (TREM2)	TREM2
Tumor Necrosis Factor Receptor-I (TNF-RI)	TNFRSF1A

Table 2: The table lists the diverse therapeutic targets catalogued in AlzPED that have been nominated by researchers from NIA's AMP-AD Consortium.

Animal Models

Within the 1402 curated studies compiled in AlzPED, 6 different animal species have been utilized, a majority of which are mouse models of AD (Figure 13). Other animal species include rat, guinea pig, rabbit, dog, and non-human primate models of AD. Preclinical efficacy data from 59 model types and 226 different AD animal models are currently available in AlzPED. Also included are data from 28 mouse models of ADRD (including tauopathies and synucleinopathies). Preclinical efficacy data from new AD animal models generated in the NIA-established Model Organism Development and Evaluation for Late-Onset Alzheimer's Disease (MODEL-AD) Consortium will be included as they become available.

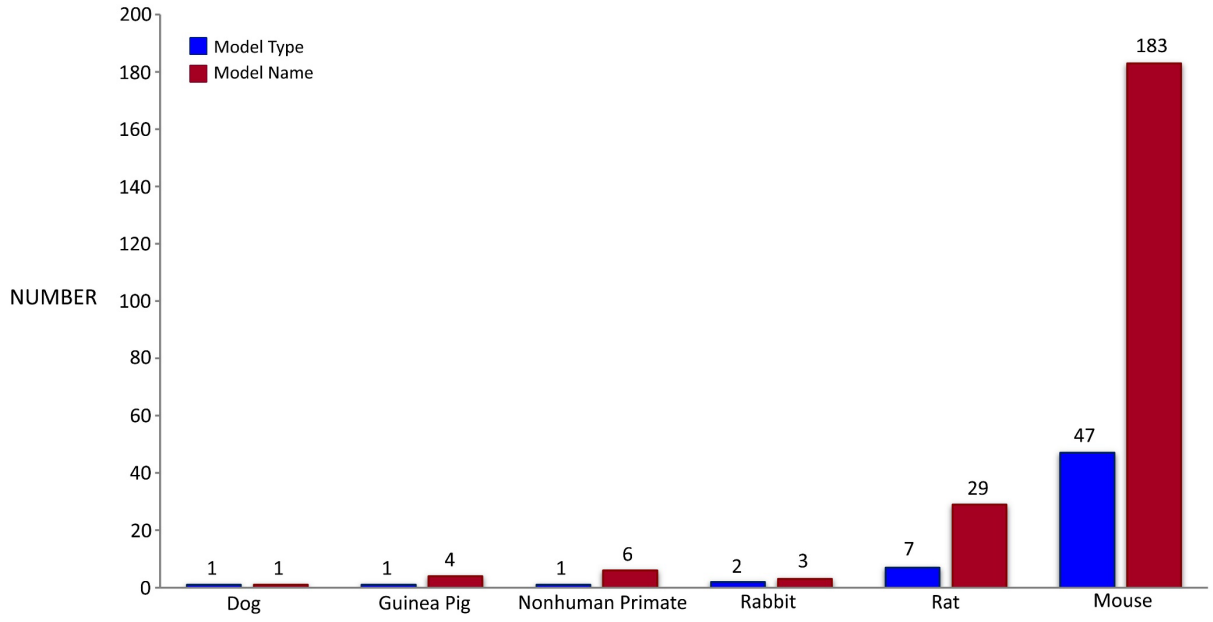

Figure 13: 6 animal species, 59 different animal model types and 226 AD animal models are utilized in preclinical efficacy studies, derived from 1402 published preclinical studies curated to AlzPED.

Outcome Measures

Each curated study provides an individual snapshot of the measures tested and outcomes achieved in response to the therapeutic agent tested. AlzPED defines 22 major outcome measures that are categorized as either functional or descriptive (Figure 14).

Functional measures include behavioral, motor, electrophysiological and imaging outcomes. Of these functional measures, behavioral outcomes are most commonly tested. There are 90 unique behavioral outcomes measured, from which the Morris water maze, novel object recognition, open field tests and Y maze are the most frequently studied. Within the 26 different motor function outcomes measured, locomotor activity, swimming speed, path length and the rotarod test are the most frequently studied. 79 diverse electrophysiological outcomes are measured, the most frequently measured being long term potentiation, field excitatory postsynaptic potentials, paired pulse facilitation and input/output (I/O) curve. Within the 55 unique imaging outcomes measured, cerebral blood flow, structural MRI and *in vivo* two-photon amyloid and calcium imaging are the most frequently studied.

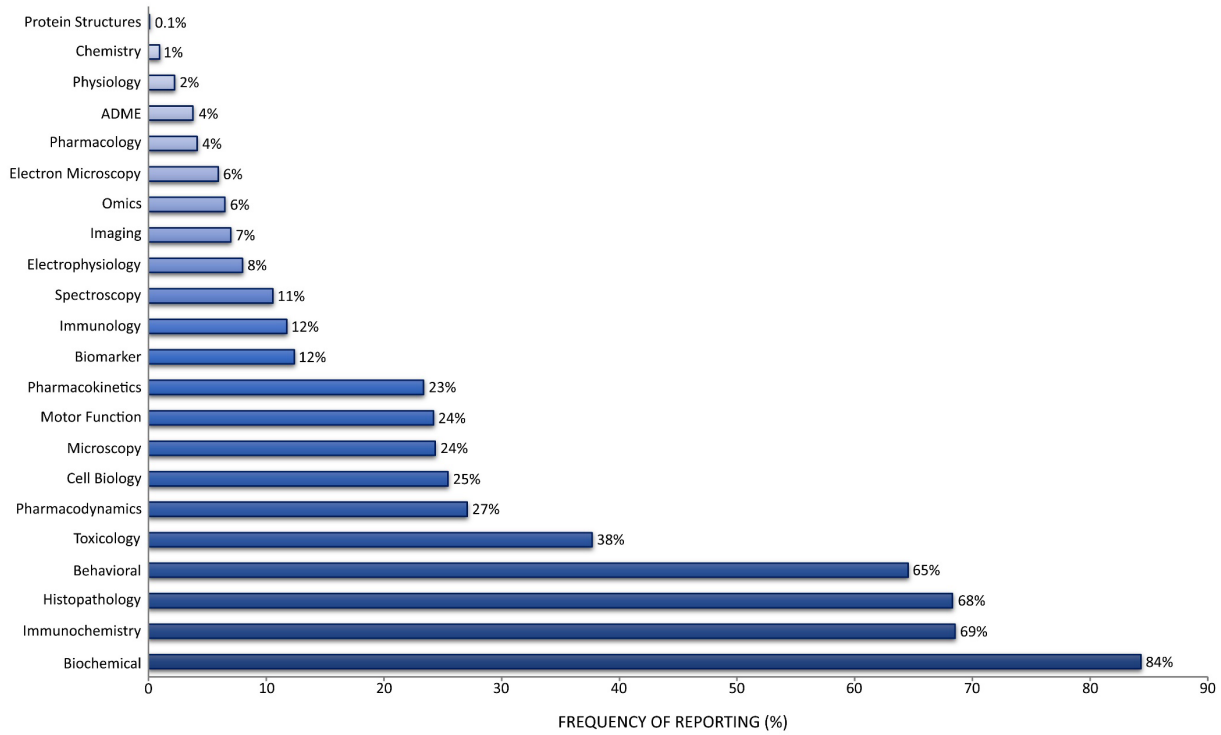

Figure 14: The graph shows the percentage of studies reporting 22 major outcome measures, derived from 1402 published preclinical studies curated in AlzPED.

Descriptive measures include ADME, biochemical, biomarker, cell biology, chemistry, electron microscopy, histopathological, immunochemical, immunological, microscopy, omics (proteomics, lipidomics, metabolomics, transcriptomics, and others), pharmacodynamic, pharmacokinetic, pharmacological, physiological, protein structures, spectroscopy, and toxicology outcomes.

Within the descriptive measures tested, beta amyloid pathology-related biochemical, histopathological, immunochemical, and biomarker outcomes are a major focus in the studies curated to AlzPED. These measures analyze several species of beta amyloid including soluble, insoluble, monomers, oligomers, fibrils, and plaques. Other measures in these categories include evaluation of several species of tau (soluble, insoluble, aggregated, hyperphosphorylated, and others), as well as astrocytic, microglial, and synaptic markers.

Notably, even though beta amyloid, tau species, and glial markers are a major focus, an extraordinary range of factors and molecules are investigated within these 4 descriptive measures. In total, information from 1558 biochemical, 40 histopathological, and 461 immunochemical measures are currently available in AlzPED. As many as 40 different biomarkers have been analyzed, and beta amyloid markers in plasma, serum, or CSF constitute a large proportion.

Other frequently studied descriptive measures used to characterize the therapeutic agent being tested include ADME, pharmacokinetic, pharmacodynamic and toxicology outcomes. Of the 30 ADME

measures studied, the most commonly tested are biodistribution, metabolic stability, and cytochrome p450 inhibition capability of the therapeutic agent. Similarly, 124 different pharmacodynamic measures are examined with key focus on reducing beta amyloid species. As many as 60 pharmacokinetic measures have been analyzed, with drug concentration in brain and plasma being evaluated most frequently. A comprehensive listing of at least 118 toxicology measures such as Ames tests, enzyme profiles, organ histology, and others are available in the database as well. Of these, the most frequently evaluated are body weight, general behavior, and food and water intake.

AlzPED reports on 20 different physiological measures from which blood pressure and cerebral blood flow are most frequently evaluated, as well as 8 pharmacological measures from which the most commonly tested are binding affinity and target selectivity of the therapeutic agent. As many as 140 cell biology outcomes are measured, of which cell viability and cytotoxicity are the most common measures. Of the 36 immunological measures reported, antibody titers and target specificity are most frequently evaluated. AlzPED also informs on 39 OMICS-related measures such as metabolomics and gene expression profiles, 44 electron microscopy outcomes, 74 microscopy outcomes and 24 spectroscopy outcomes. Finally, AlzPED also reports on protein structure and chemistry outcomes. In summary, there are more than 2500 AD-related outcomes measured in 1402 preclinical studies curated in this database.

SUMMARY

Human clinical trial designs are based on results from late-stage preclinical animal research. The success of these trials depend on the quality of the preclinical research, so there is a high cost when animal research fails to be reliable or reproducible. Poorly designed preclinical studies can lead to mistaken results that slow the progress of science, lead the scientific community astray, waste valuable resources, waste animal life, and lead to high rates of failure in the development of novel therapeutics. This is especially true for AD therapy development. Thus far, there are very few drugs approved by the FDA and these have a mild impact on symptoms in a subset of patients, have no effect on disease progression and mortality, and may produce serious adverse effects.

Rigorous experimental design and transparent reporting are clearly essential if animal studies are to inform future research, science policies, and successful clinical trials. AlzPED prioritizes and promotes the use of rigorous methodology in the planning of animal studies through its Rigor Report Card. The Rigor Report Card, through its checklist of experimental design elements, creates awareness of reporting recommendations and standards early in the research process and provides a practical and easy approach to planning a therapeutic study in animals. It also provides a starting point to evaluate the quality of reporting practices as well as assess the effectiveness of policies and interventions intended to improve the reporting of animal experiments and their predictive validity.

Ultimately, rigorous science and its comprehensive and transparent reporting are our responsibility – in how we design, execute, and report experiments, in how we conduct peer review of grants and papers, and in how we teach the next generation of scientists about rigor, reproducibility, and transparent reporting of research. Raising the bar for preclinical studies will require a concerted effort from the scientific community, funding agencies, and scientific journal publishers, so that funded and published research are sufficiently rigorous, transparent, and reproducible.