Challenges and opportunities in Neuroscience - Updates from the ASA BIOP SWG of Neuroscience

Mandy Jin (AbbVie), Jianchang Lin (Takeda), Jia Jia (AbbVie), Hui Yang (Astellas), Bo Lu (OSU), Inna Perevozskaya (BMS), Xiaodong Luo (Sanofi), Yixin Fang (AbbVie)

ASA BIOP DL Webinar March 2026

Disclaimer

The comments provided here are solely those of the presenters and are not necessarily reflective of the positions, policies or practices of their employers.

Introduction to the ASA Biop Section Scientific Working Group for Neuroscience

Mandy Jin (AbbVie)

ASA Biopharmaceutical Section Scientific Working Group for Neuroscience

- Co-chairs: Mandy Jin (AbbVie), Jianchang Lin (Takeda)
- Launched in Dec 2024
- Core members

Team 1 Statistical methods for neuroscience :
Co-lead: Jia Jia (AbbVie), Hui Yang (Astellas)

Members:
- Jianchang Lin (Takeda)
- Jonathan Hartzel (Merck)
- Yi Lu (Sanofi)
- Nan Hu (Genentech)
- Mandy Jin (AbbVie)
- Qi Qi (Genentech)

Team 2 Innovative study designs in neuroscience:
Co-lead: Bo Lu (OSU), Inna Perevozskaya (BMS)

Members:
- Shouhao Zhou (PSU)
- Mandy Jin (AbbVie)
- Kan Li (Merck)
- Shyla Jagannatha (Johnson and Johnson)
- Michelle Zhang (Stealth BioTherapeutics)

Team 3 AI/ML for neuroscience:
Co-lead: Yixin Fang (AbbVie), Xiaodong Luo (Sanofi)

Members:
- Haoyan Hu (Eli Lilly)
- Eric Zhang (Eikon Therapeutics)
- Yilong Zhang (Meta)
- Dooti Roy (Boehringer Ingelheim)
- Jianchang Lin (Takeda)
- Ye Li (FDA)

Complex and Diverse Landscape of Neurological Diseases

Examples of Diseases in Neuroscience

- Neurodegenerative Diseases: Alzheimer's Disease, Parkinson’s Disease, Huntington’s Disease, Amyotrophic Lateral Sclerosis (ALS)
- Neuropsychiatric Disorders: Depression, Anxiety Disorders, Schizophrenia
- Neurological Disorders: Epilepsy, Stroke, Multiple Sclerosis
- Movement Disorders: Essential Tremor, Dystonia
- Pain Syndromes: Neuropathic Pain, Migraine

Challenges for neuroscience clinical trials

Key Challenges:

- Complex biological pathologies
- Subjective outcome measures
- High placebo effects
- Patient compliance and missing data
- Complex and variable data patterns
- High costs and lengthy development timelines
- Barriers to innovative trial designs
- More

Goals of ASA BIOP NS SWG

- Acquire and advocate advanced statistical methodology for neuroscience clinical development
- Bridge gaps in the guidance of neuroscience clinical trials by engaging cross-company statisticians and regulatory workers, and to facilitate utilization and implementation of innovative approaches in clinical research for neuroscience diseases
- Share knowledge and provide education through presentations, conferences, and workshops
- Publish case studies and white papers on key neuroscience topics in peer-reviewed journals

Goals of the Teams

Team 1. Statistical methods for neuroscience

- Explore challenging, disease-specific problems in neuroscience
- Develop a library of methods for such problems
- Share best practices and lessons learned

Team 2. Innovative designs for neuroscience

- Evaluate clinical and regulatory contexts to identify optimal trial designs
- Review regulatory guidance for various designs and highlight gaps in implementation
- Develop points to consider for study design that is feasible, efficient, and interpretable

Team 3. AI/ML in neuroscience

- Research and understand existing methodologies of AI/ML
- Assess the opportunities to utilize AI/ML in neuroscience
- Share findings in conferences or journals

Updates in Team 1: Statistical methods in Neuroscience

Jia Jia* (AbbVie), Hui Yang (Astellas)

Team 1. Statistical methods for neuroscience

- Explore challenging, disease-specific problems in neuroscience
  - Placebo Effect
  - Subject outcome measures
  - Missing data in Neuroscience
  - Small signal-to-noise ratio
- Evaluate methods for such problems
  - Traditional Statistical Methodologies
  - Recent development based on the traditional methodologies, in specific to analyze data from Neuroscience RCTs.
  - Prognostic Covariate Adjustment (PROCOVA)
- Share best practices and lessons learned
  - Gather examples from RCTs across industry.

Member

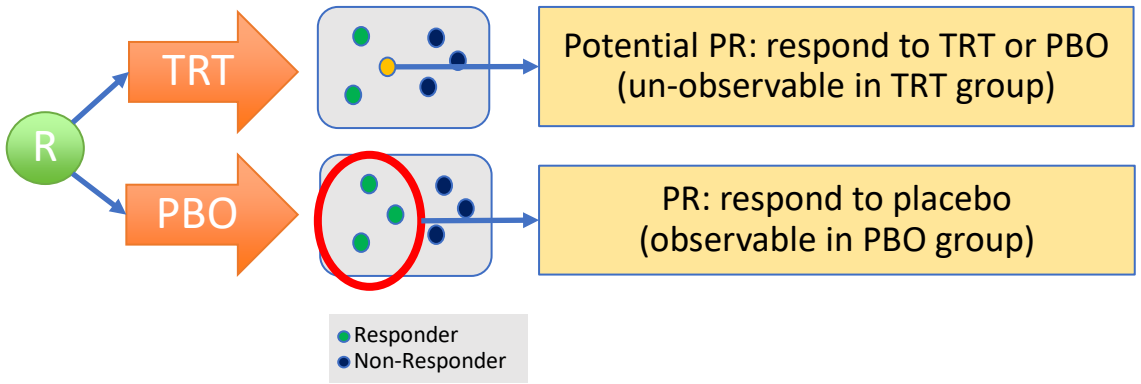
- Jianchang Lin (Takeda)
- Jonathan Hartzel (Merck)
- Yi Lu (Sanofi)
- Nan Hu (Genentech)
- Mandy Jin (AbbVie)
- Qi Qi (Genentech)
- Jia Jia (AbbVie)
- Hui Yang (Astellas)

Placebo Effect

Patients in randomized controlled trials (RCT) can experience real improvements from a placebo due to the **placebo effect**, which is driven by their expectations, and is impacted by many factors.

Especially impacts subjective outcomes which are commonly implemented in Neuroscience RCTs. For example:

- Pain scores
- Depression rating scales



- Placebo Effect impacts patients in both TRT and PBO groups, but only partially observable.
- A 2016 Lancet review found the average placebo response rate in antidepressant trials is about **35–40%**[1].
- Many clinical trials are failed due to high placebo effect.

Complex Designs and an Innovative Weighted Method

Innovative design options are available, but with limitations.

Gomeni et al 2023[2] paper proposed a machine learning approach combined with prognostic score weighting of probability of being a placebo responder to control placebo effect in depressive disorder.

Limitations[3]:

- Repeated data usage, weights built from the in-trial postbaseline and baseline data (Rogers and Senn 2025[2])

We propose to build the prediction model based on **historical data (avoid repeated use of the postbaseline data in the current trial)**, then create weights for the current RCT using baseline data.

ML model powered weights + Weighted analysis
(Methodology: Estimand: Weighted ATE, Unbiased estimators)

Conducted simulation and used real data example to evaluate this weighted methods powered by Machine Learning models.

Our work has shown great promising of this approach. A manuscript[4] summarizing the work has been submitted.

ML: machine learning; ATE: averaged treatment effect

Subjective Outcome Measurements

Subjectivity and patient context in outcome measures complicate comparison and reliability of treatment effects in neuroscience trials.

- Subjective outcome measures in neuroscience
  - Patient-reported: Rely on individual experience and perception
  - Rated by investigators: not fully objective.
- Other factors can also impact outcome measures:
  - Study design
  - Logistics
  - Patient experience (learning effect)
- Two studies on progressive muscle weakness:
  - ICE[5]: traditional design
  - PATH[6]: Randomized withdrawal design

Different outcomes (change from baseline) on CIDP patients
- ICE: positive trend
- PATH: negative trend

Patients tend to be **more optimistic** when receiving a new treatment but **expect only stability** (no worsening) if moved to standard care.

Such variability complicates the comparability of results between and within studies

CIDP: Chronic Inflammatory Demyelinating Polyradiculoneuropathy

Missing Data In Neuroscience Clinical Trials

Neuroscience clinical trials often face higher levels of missing data (e.g., up to 37%[7] missing in active treatment group).

- Subjective outcome measures
- Missed visits.
- Treatment discontinuation.

Missing at Random (MAR) assumption

- MMRM/GLMM with multiple imputation (MI) are valid.
- Commonly used for efficacy analysis.

Missing Not at Random (MNAR) assumption

- Usually no evidence to justify MAR assumption.
- Sensitivity analysis to assess the robustness of the analysis results under MAR.

Imputation methods under **MNAR** assumption

- **Reference-based imputation (RBI):**
  - Jump to Reference (J2R): assumes immediate loss of treatment benefit upon discontinuation (most conservative).
  - Copy Reference (CR): allows gradual transition to reference group effects.
  - Copy Increments in Reference (CIR): matches incremental changes post-discontinuation (least conservative, hardest to justify).
- **Return to Baseline (RTB) imputation**
  - Imputes missing data using baseline values.
  - May overestimate recovered measurements in progressive neurodegenerative diseases (e.g., Alzheimer's)

Lower than expected Signal-to-Noise ratio

- Continuous endpoints (e.g., change from baseline) are standard as primary endpoints in neuroscience Phase 3 trials.
- Accurate estimation of signal-to-noise ratio is critical for study power, but placebo effects and unexpected patient trajectories can reduce this ratio.
- Some trials have observed placebo groups not declining as expected, complicating endpoint interpretation.
- Inadequate endpoint validation and changing patient populations can further increase uncertainty.
- Adaptive and Type I error controlled methods are being explored to improve endpoint selection.[19, 20]
- Better models of natural disease progression and translation between early and late-stage endpoints are needed for optimized trial design.

Evaluation of Longitudinal Statistical Models

Mixed Model for Repeated Measures (MMRM)[8]

- Common in RCTs; valid under MAR; visit-specific effects.

Constrained Longitudinal Data Analysis (cLDA)[9]

- Baseline as outcome; similar baseline means across arms.

Progression Models for Repeated Measures (PMRMs)[10]

- Continuous-time mixed-effects progression models.

Natural Cubic Splines (NCS)[11]

- Smooth flexible time basis within mixed model.

Random Coefficient Regression Models (RCRM)[12]

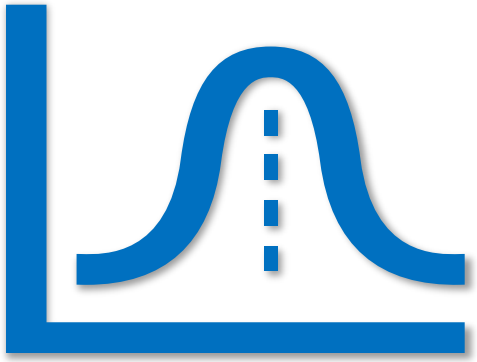
- Allow individual variability in time trajectories.

Proportional cLDA (PcLDA) and pMMRM[13]

- Proportional treatment effect over time.

Bayesian Disease Progression Models (DPMs)[14, 15]

- Hierarchical nonlinear progression aligned by (latent) time since onset.



No single model suits all neuroscience trials; choice depends on disease, trial design, dropout, regulatory context.

MMRM and cLDA are foundational due to robustness and acceptance.

Spline, proportional, and progression models offer further advantages; hybrid/flexible modeling may enhance future trials.

Evaluation of Analysis Approach - Prognostic Covariate Adjustment (PROCOVA)

How It Works:

1. Gather relevant **historical/external data**.
2. Develop a prediction model using baseline covariates to **estimate control outcomes**.
3. Apply this model to calculate **prognostic scores** by predicting potential outcomes for current trial participants.
4. Analyze trial results using a linear model while adjusting for the **prognostic score** as a covariate

Pros: Can improve the precision of estimated treatment effects.

Cons: Depends on quality and relevance of historical data and the performance of prediction model.

Within a defined context of use, PROCOVA is accepted by EMA's qualification opinions for Prognostic Covariate Adjustment (PROCOVA™) and ICH E9(R1) (estimands and sensitivity analyses) : **adjustment is encouraged when it improves precision and is fully pre-specified**.

European Medicines Agency (EMA), Committee for Medicinal Products for Human Use (CHMP). 2023. Qualification opinion: Prognostic Covariate Adjustment (PROCOVA™) EMADOC-1700519818-907465.

References

1: Furukawa, Toshi A., et al. "Placebo response rates in antidepressant trials: a systematic review of published and unpublished double-blind randomised controlled studies." The Lancet Psychiatry 3.11 (2016): 1059-1066.

2: Gomeni, Roberto, et al. "Interpreting clinical trial outcomes complicated by placebo response with an assessment of false-negative and true-negative clinical trials in depression using propensity-weighting." Translational Psychiatry 13.1 (2023): 388.

3: Rogers, J.A., Senn, S. Randomization and placebo effects in clinical trials of major depressive disorder. Transl Psychiatry 15, 43 (2025). https://doi.org/10.1038/s41398-025-03263-0

4: Jia Jia, Diezhang Wu, Man Jin (2026). Statistical Methods to Control Confounding Placebo Effect in Randomized Clinical Trials Using Machine Learning Methods. Journal of Biopharmaceutical Statistics

5: Hughes RA, Donofrio P, Bril V, Dalakas MC, Deng C, Hanna K, Hartung HP, Latov N, Merkies IS, van Doorn PA. Intravenous immune globulin (10% caprylate-chromatography purified) for the treatment of chronic inflammatory demyelinating polyradiculoneuropathy (ICE study): a randomised placebo-controlled trial. The Lancet Neurology. 2008 Feb 1;7(2):136-44.

6: van Schaik IN, Bril V, van Geloven N, Hartung HP, Lewis RA, Sobue G, Lawo JP, Praus M, Mielke O, Durn BL, Cornblath DR. Subcutaneous immunoglobulin for maintenance treatment in chronic inflammatory demyelinating polyneuropathy (PATH): a randomised, double-blind, placebo-controlled, phase 3 trial. The Lancet Neurology. 2018 Jan 1;17(1):35-46.

7: Michael J Soileau, et al. Safety and efficacy of continuous subcutaneous foslevodopa-foscarbidopa in patients with advanced Parkinson's disease: a randomised, double-blind, active-controlled, phase 3 trial. The Lancet Neurology, Volume 21, Issue 12, 2022, Pages 1099-1109.

8: Laird, n.M and Ware, J.h. Random-effects Models for Longitudinal Data. Biometrics, 38 (1982), 963–974.

9: Liang KY, Zeger SL. Longitudinal data analysis of continuous and discrete responses for pre-post designs. Sankhy a: Indian J Stat Ser B.2000;62(1):134-148.

10: Raket, L.L. (2022), A new approach for estimating and quantifying disease-modification in Alzheimer’s disease clinical trials. Alzheimer's Dement., 18: e063092.

11: Donohue MC, Langford O, Insel PS, et al. Natural cubic splines for the analysis of Alzheimer's clinical trials. Pharmaceutical Statistics. 2023;22(3):508-519. doi:10.1002/pst.2285

12: Hu N, Mackey H, Thomas R. Power and sample size for random coefficient regression models in randomized experiments with monotone missing data. Biom J. 2021 Apr;63(4):806-824. doi: 10.1002/bimj.202000184. Epub 2021 Feb 15.

13: Wang G, Liu L, Li Y, et al. Proportional constrained longitudinal data analysis models for clinical trials in sporadic Alzheimer's disease. Alzheimer's & Demen: Transl Res Clin Interv. 2022;8(1):e12286.

14: Wang, Guoqiao et al. “A novel cognitive disease progression model for clinical trials in autosomal-dominant Alzheimer's disease.” Statistics in medicine vol. 37,21 (2018): 3047-3055. doi:10.1002/sim.7811

15: Staffaroni, Adam M et al. “Temporal order of clinical and biomarker changes in familial frontotemporal dementia.” Nature medicine vol. 28,10 (2022): 2194-2206. doi:10.1038/s41591-022-01942-9

16: Cro S, Morris TP, Kenward MG, Carpenter JR. Reference-based sensitivity analysis via multiple imputation for longitudinal trials with protocol deviation. Stata J. 2016 Apr;16(2):443-463. PMID: 29398978; PMCID: PMC5796638.

17: Liu GF, Pang L. On analysis of longitudinal clinical trials with missing data using reference-based imputation. J Biopharm Stat. 2016;26(5):924-36. doi: 10.1080/10543406.2015.1094810. Epub 2015 Sep 29. PMID: 26418282.

18: Schuler A, Walsh D, Hall D, Walsh J, Fisher C; Critical Path for Alzheimer’s Disease; Alzheimer’s Disease Neuroimaging Initiative; Alzheimer’s Disease Cooperative Study. Increasing the efficiency of randomized trial estimates via linear adjustment for a prognostic score. Int J Biostat. 2021 Dec 22;18(2):329-356. doi: 10.1515/ijb-2021-0072. PMID: 34957728.

19: Nan Hu, Godwin Yung & Howard Mackey (12 Mar 2025): Type I Error Control of an Adaptive Endpoint Selection Procedure, Statistics in Biopharmaceutical Research, DOI:10.1080/19466315.2025.2455173

20: Xu, H., Liu, Y., & Beckman, R. A. (2023). Adaptive Endpoints Selection with Application in Rare Disease. Statistics in Biopharmaceutical Research, 16(1), 55–63. https://doi.org/10.1080/19466315.2023.2183252

Updates in Team 2: Innovative study designs in Neuroscience

Bo Lu* (OSU), Inna Perevozskaya (BMS)

Team members

(in alphabetical order)

Shyla Jagannatha (Johnson & Johnson)

Mandy Jin (AbbVie)

Kan Li (Merck)

Bo Lu (Ohio State University)

Inna Perevozskaya (BMS)

Michelle Zhang (Stealth BioTherapeutics)

Shouhao Zhou (Pennsylvania State University)

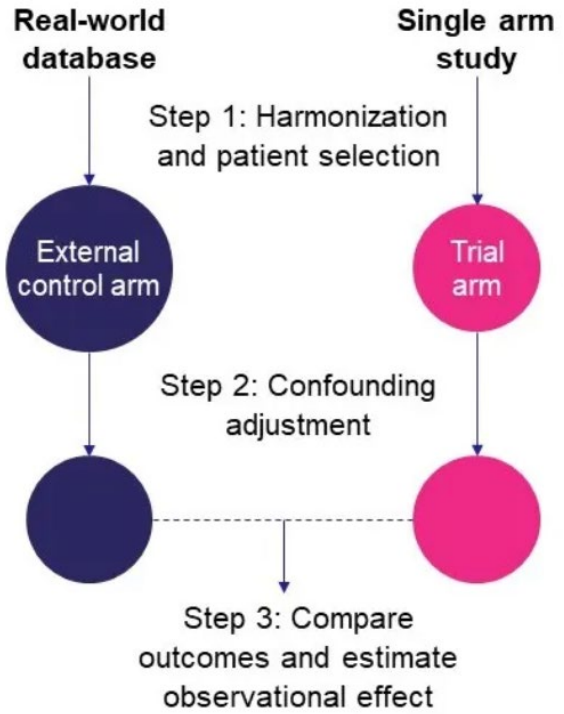
Challenges and Opportunities in Trial Designs

- Neuroscience therapeutic areas face unique challenges: low prevalence in some neurologic disorders, slow progression, high placebo effect, long-term follow-up, high variability in patient responses, etc.
- We explore four topics:
  - External control borrowing
  - Platform trial
  - Adaptive design
  - Placebo effect

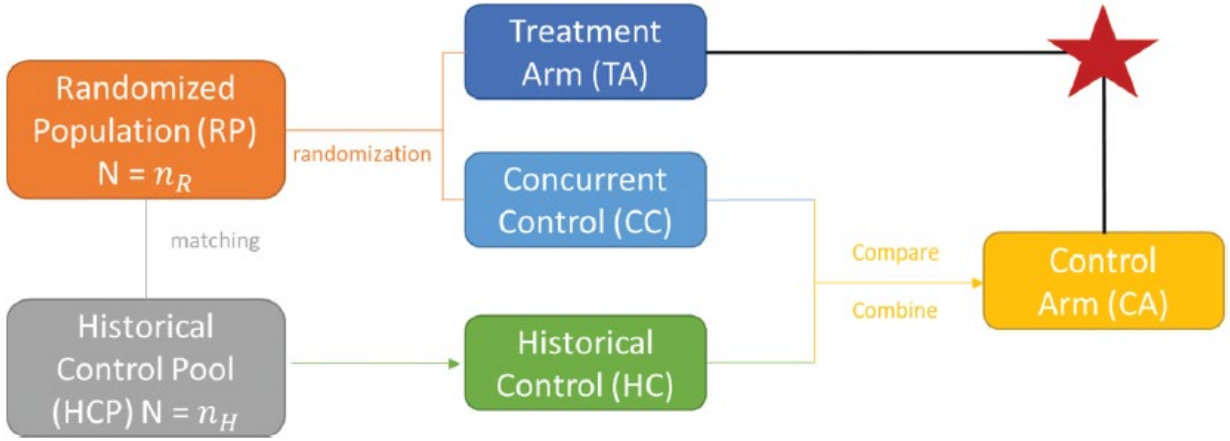
Design to leveraging real-world data

- One major limiting factor of running conventional randomized clinical trials for neurological diseases is the difficulties in recruitment due to rarity of the disease. For example, 2 cases per 100, 000 in major Western countries for ALS.
- Borrowing external control data to supplement trials (either hybrid or single-arm design) is adopted under certain assumptions.
- Two types of common approaches:
  - Bayesian borrowing methods
  - Propensity score-based methods

Single arm vs. Hybrid designs



https://cytel.com/perspectives/the-role-of-external-data-in-oncology-drug-development/



Liu, Lu, et al. (2022) Matching design for augmenting the control arm of a randomized controlled trial using real-world data, Journal of Biopharmaceutical Statistics, 32:1, 124-140

Bayesian methods

- Key idea: Incorporate external data information into the prior distribution for the causal parameter and modulate the prior contribution based on the similarity with trial data.
- Common methods
  - Power prior
  - Meta-analytic prior (MAP)
  - Commensurate prior
- Limitations

Discount external data information at equal rate (not distinguishing between good and bad controls); performance may be sensitive to the choice of priors.

- One example of Bayesian: Bayesian pediatric design (Bovis F, Ponzano M, Signori A, Schiavetti I, Bruzzi P, Sormani MP. Reinterpreting Clinical Trials in Children With Multiple Sclerosis Using a Bayesian Approach. JAMA Neurol. 2022;79(8):821–822. doi:10.1001/jamaneurol.2022.1735)

Propensity score methods

- Key idea: using propensity score-based design tools to remove confounding bias from the external data.
- Common methods
  - Propensity score matching
  - Propensity score weighting
  - Propensity score stratification
- Limitations

Depend on external data quality; Rely on exchangeability assumptions, which may require sensitivity analysis.

Platform trials

- Traditional trial designs are inefficient and expensive for neurological diseases, due to slow progression, long follow-up, multiple mechanisms of disease.
- A platform trial is designed to evaluate multiple treatments within a single, ongoing master protocol, integrating adaptive randomization, allowing treatments to be added or dropped over time.
- Unlike a typical multi-arm trial conducted only once, a platform design is “perpetual” or open-ended: as new therapeutic candidates emerge, they can be integrated into the ongoing infrastructure.
- The shared-control design reduces the need for repetitive control arms, and attractive to patients with unmet medical needs such as ALS or Alzheimer’s.

Platform trials-Cont'd

- Due to multiple and complex decisions to be made (early stopping, changing allocation ratio etc.), the statistical framework often utilizes Bayesian designs with interim analyses that may include borrowing of information across arms and longitudinal modelling across time.
- Platform trial designs also unlock several key opportunities for neuroscience research:
  - The ability to integrate external or historical controls into the shared control arm
  - Efficient algorithms of adaptive randomization and early stopping rules can enable assignment weights to shift toward more promising arms
  - The synergy with emerging digital biomarkers, wearable data, and AI–driven interim analyses
  - Can facilitate seamless phase II/III transitions
- One example of Platform Trial: ALS platform design (Paganoni S, Berry JD, Quintana M, Macklin E, Saville BR, Detry MA, Chase M, Sherman AV, Yu H, Drake K, Andrews J, Shefner J, Chibnik LB, Vestrucci M, Cudkowicz ME; Healey ALS Platform Trial Study Group. Adaptive Platform Trials to Transform Amyotrophic Lateral Sclerosis Therapy Development. Ann Neurol. 2022 Feb;91(2):165-175. doi: 10.1002/ana.26285. Epub 2022 Jan 10. PMID: 34935174.)

Adaptive design

- Adaptive designs allow pre-specified modifications to key trial elements—such as sample size, randomization ratios, dosing, or treatment arms—based on interim data analyses while preserving statistical validity and trial integrity.
- In neurological diseases, patient heterogeneity, uncertain effect sizes, slow disease progression, and limited eligible populations are common challenges, where adaptive designs can improve efficiency and ethical conduct.
- Sample size re-estimation can better address uncertainty in effect size or variance.
- A key challenge in implementation of adaptive designs in neuroscience is endpoint availability.

Placebo effects

- Placebo effect refers to the improvement in symptoms that occurs simply from receiving a substance or undergoing a procedure, even when the substance or procedure has no active therapeutic effects.
- The sources of placebo effects are complex and multifactorial, and expectancy-based mechanisms play an important role in placebo effects.
- Mitigating the placebo effect in RCTs is challenging:
  - The treatment group may contain unobserved “potential placebo responders”
  - Complex data structure including baseline covariates, imaging data, biomarkers, and other factors may influence treatment response and can be difficult to control
  - Regulatory agencies often limit the number of covariates that can be included in an analysis

Designs to mitigating placebo effects

- Placebo lead-in design

  All participants receive placebo during a run-in phase before randomization. Those showing early improvement are then excluded to remove individuals most susceptible to placebo effects before randomization.

- Randomized-withdrawn design

  All participants receive the active treatment during a pre-randomization period. Responders are then randomized to either continue active treatment or switch to placebo. Those showing worsened outcomes after switching show true treatment effect.

- Sequential parallel comparison design (SPCD)

  In Phase 1, all participants are randomized to drug or placebo. In Phase 2, placebo non-responders from Phase 1 are re-randomized to drug or placebo, while drug responders continue drug. It enriches Phase 2 with placebo non-responders, improving signal detection.

References

- ALS platform design: Paganoni S, Berry JD, Quintana M, Macklin E, Saville BR, Detry MA, Chase M, Sherman AV, Yu H, Drake K, Andrews J, Shefner J, Chibnik LB, Vestrucci M, Cudkowicz ME; Healey ALS Platform Trial Study Group. Adaptive Platform Trials to Transform Amyotrophic Lateral Sclerosis Therapy Development. Ann Neurol. 2022 Feb;91(2):165-175. doi: 10.1002/ana.26285. Epub 2022 Jan 10. PMID: 34935174.

- Bayesian pediatric design: Bovis F, Ponzano M, Signori A, Schiavetti I, Bruzzi P, Sormani MP. Reinterpreting Clinical Trials in Children With Multiple Sclerosis Using a Bayesian Approach. JAMA Neurol. 2022;79(8):821–822. doi:10.1001/jamaneurol.2022.1735

Updates in Team 3: AI/ML for Neuroscience

Xiaodong Luo* (Sanofi), Yixin Fang (AbbVie)

Why it worked?
How it worked?
Can it be translated to my case?

Note: this image was generated by Chatgpt, following a command “generate an image that shows ‘all you need is to learn and understand a use-case example’ in applying AI/ML in neuroscience”.

AI/ML for Neuroscience

- Regulatory landscape of AI/ML in drug development
  - 6 examples
  - 7-step risk-based credibility assessment framework
- AI/ML in forming a sound research question-PICOT(S)
  - 6 components
- AI/ML in neuroscience clinical trials
  - 5 stages

Members
- Yixin Fang (AbbVie)
- Haoyan Hu (Eli Lilly)
- Jianchang Lin (Takeda)
- Xiaodong Luo (Sanofi)
- Dooti Roy (Boehringer Ingelheim)
- Eric Zhang (Eikon Therapeutics)
- Yirong Zhang (Meta)

Regulatory landscape of AI/ML in drug development

In January 2025, FDA draft guidance for industry and other interested parties titled "Considerations for the Use of Artificial Intelligence To Support Regulatory Decision-making for Drug and Biological Products" (Jan 2025)

- Provides recommendations to sponsors and other interested parties on the use of artificial intelligence (AI) to produce information or data intended to support regulatory decision-making regarding safety, effectiveness, or quality for drugs.

- Provides 6 examples six general examples of AI applications across all therapeutic areas.

  We describe 6 specific examples of AI applications in neuroscience throughout the life cycle of drug products

- Provide a 7-step risk-based credibility assessment framework

  We will provide a skeleton example in neuroscience

6 examples

1. Preclinical: Using AI to reduce the number of animal-based PK, PD and Tox studies in neuroscience
   - AI models can analyze large datasets from *in vitro* and *in silico* experiments to predict PK/PD profiles[1], blood-brain barrier penetration[2], and toxicity of new compounds[3]
2. Early phase trials: Using predictive modeling for clinical pharmacokinetics and/or exposure-response analyses in neuroscience.
   - In AD, hybrid ML–PK/PD frameworks predict central exposure and clinical outcomes (eg.,CDR-SB, ADAS-Cog) from plasma PK, biomarker (Aβ, tau), and imaging data[4]
3. Understanding the diseases: Integrating data from various sources to improve understanding of neurological disease presentations, heterogeneity, predictors of progression, recognition of disease subtypes.
   - Gene-SGAN[5], a multi-view, weakly-supervised deep clustering method designed to dissect disease heterogeneity by jointly considering phenotypic and genetic data
   - Smile-GANs[6], a semi-supervised clustering method that utilizes generative adversarial networks to analyze brain MRI
   - MAGIC[7]: an algorithm employing multi-scale clustering to uncover disease heterogeneity in AD
   - Trajectory Clustering (TC)[8]: a network-based algorithm to identify PD subtypes based on disease trajectory by modeling patient-variable interactions as a bipartite network,
4. Late phase trials: Processing and analyzing large sets of data for the development of clinical trial endpoints or assessment of outcomes for neurological diseases.
   - Leveraging RWD and AI in neurology[9]: used longitudinal RWD to emulate clinical trials and identify potential disease-modifying therapies for PD.
   - Deep learning for low-count PET enhancement: techniques can process low-count PET images to produce diagnostic-quality outputs comparable to full-dose or full-duration acquisitions[10].
5. Postmarketing: Identifying, evaluating, and processing for reporting postmarketing adverse drug experience information in neuroscience.
   - A Korean study utilized supervised ML algorithms to analyze adverse event data from the Korea Adverse Event Reporting System (KAERS) to detect patterns indicative of potential safety signals[11]
6. Manufacturing: Facilitating the selection of manufacturing conditions.
   - TBD

Deep Learning for Low-Count PET Enhancement

Core Idea

- Evaluated a DL algorithm to enhance **low-count PET scans** (~25% of normal counts → 4× lower dose/scan time).
- Enhanced images were compared to **full-dose PET** across **multicenter, multivendor datasets**.

Key Findings

- Diagnostic quality (DIQ) & confidence (ODC): Noninferior to full-dose PET (DIQ $p < 0.05$; ODC $p < 0.001$)
- Lesion detection: Sensitivity ~0.94, Specificity ~0.98, Kappa ~0.85
- Quantitative accuracy: SUV correlation ≥0.94 with standard PET

How It Worked

- 2.5D encoder-decoder CNN (SubtlePET®)
- Trained on separate datasets; externally validated across multiple scanners and institutions

Why It Matters

- Clinical utility: Maintains diagnostic quality & quantitative accuracy with ~75% fewer counts
- Benefits:
  - Reduced radiation exposure
  - Shorter scan times → better patient experience
  - Cost savings
  - Generalizable across clinical sites

AI enhances each element of PICOT(S)

- Population: AI/ML allows unprecedented analysis of heterogeneous patient populations by leveraging largescale electronic health records (EHRs), neuroimaging datasets, and genomic data.
  - *Through clustering algorithms and pattern recognition, AI helps define subpopulations (e.g., Alzheimer's patients with specific genetic risk profiles) with greater precision, supporting targeted hypotheses and stratified study designs*[12,13].
- Intervention: AI/ML technologies play a pivotal role in identifying, classifying, and simulating interventions within neuroscience research.
  - *Reinforcement learning algorithms have been employed to optimize the simulations prior to clinical implementation*[14].
  - *Natural language processing (NLP) techniques are increasingly used to extract nuanced intervention details from biomedical literature, thereby enriching meta-analyses and informing protocol development*[15].
- Comparator: AI/ML can match cohorts or simulate control groups
  - *Digital Twins in Multiple Sclerosis Clinical Trials*[16].
- Outcome: ML-based models help define, detect, and forecast outcomes with greater nuance, from cognitive decline trajectories to quantitative changes on neuroimaging.
  - *Predictive analytics allow continuous monitoring and adaptive outcome refinement*[17].
  - *AI-COA™: a drug development tool intended to improve the reliability of depression and anxiety severity measurement while preserving established interview-based endpoints*[18]
- Time: Recurrent neural networks and survival models can uncover patterns in disease progression or treatment response.
  - *These tools inform the optimal timing for intervention delivery or outcome assessment*[19,20]
  - *Real-time AI/ML monitoring approaches in neuroscience and neurology*[21]
- Setting: AI/ML methodologies have revolutionized the design and execution of clinical trials by enabling research to extend beyond traditional clinical environments into real-world settings.
  - *The advent of decentralized clinical trials (DCTs), supported by digital health platforms and remote monitoring technologies, allows participants to engage in study visits remotely, thereby reducing the burden of travel and mitigating attrition due to illness or logistical challenges*[22-24]
  - *Remote, AI-enabled trial visit management enhancing both feasibility and generalizability of neuroscience research*[24].

AI-Generated Digital Twins for Multiple Sclerosis Trials

Core Idea

- Use **AI/ML (Conditional Restricted Boltzmann Machine, CRBM)** to generate **digital twins**: synthetic MS patient profiles that **mirror real baseline characteristics and disease progression**.
- Digital twins can act as **virtual control subjects**, supporting statistical analyses when a true control arm is small or absent.

Methodology

- Trained on **historical MS clinical trial data** (placebo arms).
- Generated **individualized synthetic patient trajectories** for disease endpoints.
- Evaluated whether digital twins were **statistically indistinguishable** from real patients.

Key Findings

- Digital twins reproduced **baseline distributions and longitudinal outcomes** with high fidelity.
- Enabled **simulated control cohorts** for trial analysis.
- Proof-of-concept shows potential to **enhance trial efficiency** and **reduce required sample sizes**.

Why It Matters

- Supports **synthetic control arms** in neurological clinical trials.
- Reduces reliance on large placebo groups, **accelerating trials**.
- Demonstrates how **AI/ML can model disease heterogeneity** at the patient level.

5 stages: AI/ML in neuroscience clinical trials

1. Planning:
   - Review existing literature to identify unmet needs; Explore preliminary data; Define target patient population; Specify inclusion and exclusion criteria; Optimize resource allocation; Forecast timeline
2. Design:
   - Define endpoints; Sample Size and Effect Size Estimation; Margin of Error Determination; Innovative Study Design Selection; Disease progression prediction
3. Conduct:
   - Real-Time Monitoring of Patient Enrollment; Safety Monitoring; Real-Time Cleaning of Data; Enabling Remote Clinical Trial Conduct; Tables, figures, and listings (TFLs) and SDTM/ADaM Dataset Generation
4. Analysis:
   - Estimation: causal inference
5. Interpretation:
   - Multimodal Data Integration; Advanced Subgroup and Responder Analysis; Biomarker Discovery and Surrogate Endpoint Validation; Sensitivity Analysis for Missing Data and Intercurrent Events Handling; Interpretability of Complex Outcome Patterns; Interpretability of Causal Inference Findings

7-step risk-based credibility assessment framework

Step 1: Define the Question of Interest

Step 2: Define the Context of Use (COU)

Step 3: Assess Model Risk

Step 4: Develop a Credibility Assessment Plan

Step 5: Execute the Plan

Step 6: Document Results

Step 7: Determine Adequacy for COU

A skeleton example

Step 1: Define the Question of Interest

- Can an AI/ML model predict individual patient cognitive outcomes (e.g., CDR-SB or ADAS-Cog change) based on early plasma PK, biomarker (Abeta, tau) data, and patient covariates? The goal is to support dose selection in early-phase AD trials

Step 2: Define the Context of Use (COU):

- The model will be used in Phase II dose-finding studies to: Predict CNS exposure and response. Identify optimal dosing regimens for patients with different baseline characteristics. Inform decisions on advancing doses to Phase III. It is not used for clinical diagnosis or treatment outside the trial.

A skeleton example (cont.)

Step 3: Assess Model Risk:

- Decision consequence: High, since model outputs influence dose selection, which affects efficacy and safety.
- Model influence: Moderate to high; model predictions support but do not fully replace traditional PK/PD analyses.
- Thus, the model is medium-to-high risk, requiring thorough validation.

Step 4: Develop a Credibility Assessment Plan. The plan includes:

- Data partitioning: training, validation, and external test sets.
- Performance metrics: RMSE for cognitive score prediction, bias, and coverage probability.
- Sensitivity analyses: evaluate robustness to missing PK/biomarker values and patient heterogeneity. Comparison with existing NLME-based exposure–response models.

A skeleton example (cont.)

Step 5: Execute the Plan. Train an ML model (e.g., XGBoost or RNN) on early-phase trial data.

- Validate against withheld data and external datasets.
- Perform stress testing to assess predictions under varying input conditions.

Step 6: Document Results. Report model performance metrics, feature importance, and limitations.

- Document data sources and data cleaning process.
- Document deviations from the plan (e.g., imputation strategies for missing CSF data).
- Include visualizations of predicted vs. observed cognitive outcomes.

Step 7: Determine Adequacy for COU

- If model performance meets predefined thresholds (e.g., RMSE, clinically meaningful change, coverage), it can be used to support dosing decisions.
- If not, additional data collection or model refinement is required.

References

[1]Oprea, et al. (2019). *Computational approaches for drug discovery and development.* Drug Discovery Today

[2] Kramer, et al. (2020).*Computational models for blood-brain barrier permeability.* Current Topics in Medicinal Chemistry

[3] Van Norman (2019). *Limitations of animal studies for predicting toxicity in clinical trials: Is it time to rethink our current approach?* JACC: Basic to Translational Science

[4] Zhang, et al. (2022).*AI-Assisted Exposure–Response Modeling for Cognitive Endpoints in Alzheimer's Disease Trials.* Clinical Pharmacology & Therapeutics

[5] Yang, et al. (2024). G*ene-SGAN: discovering disease subtypes with imaging and genetic signatures via multi-view weakly-supervised deep clustering.* Nature Communications

[6] Yang, et al. and the SMILE-GAN Consortium (2021). *Smile-GANs: Clustering Alzheimer's disease subtypes using MRI and longitudinal data.* Nature Communications

[7] Zhang, et al. (2020). *MAGIC: Multi-scale clustering for Alzheimer's disease subtypes.* Nature Communications

[8] Krishnagopal and von Coelln (2020). *Identifying and predicting Parkinson's disease subtypes through trajectory clustering via bipartite networks.* Biomedical Physics & Engineering Express

[9] Laifenfeld, et al. (2021). *Emulated clinical trials from longitudinal real-world data identify drug repurposing candidates for disease-modifying effects in Parkinson's disease.* Frontiers in Pharmacology

[10] Akshay, et al. (2021). *Low-count whole-body PET with deep learning in a multicenter and externally validated study.* npj Digital Medicine

[11] Lee, et al. (2022). *Detecting early safety signals of infliximab using machine learning algorithms in the Korea Adverse Event Reporting System.* Scientific Reports

[12]Miotto, et al. (2018). Deep learning for healthcare: review, opportunities and challenges. Briefings in bioinformatics.

References (cont.)

[13]Woo, et al. (2017). Building better biomarkers: brain models in translational neuroimaging. Nature neuroscience.

[14]Chandrabhatla , et al. (2023). Landscape and future directions of machine learning applications in closed-loop brain stimulation. NPJ Digital Medicine.

[15]Ofori-Boateng, et al. (2024). Towards the automation of systematic reviews using natural language processing, machine learning, and deep learning: a comprehensive review. Artificial intelligence review.

[16] Walsh, et al. (2020). Generating Digital Twins with Multiple Sclerosis Using Probabilistic Neural Networks. arXiv. DOI: 10.48550/arXiv.2002.02779.

[17] Esteva, et al. (2019). A guide to deep learning in healthcare. Nature medicine.

[18] Deliberate Solutions, Inc. (2025). AI-COA™ Drug Development Tool (DDT) Project. FDA Drug Development Tool (DDT) Qualification Program. 2025. Accessed December 30, 2025.

[19] Nguyen, et al. (2018). Modeling Alzheimer's disease progression using deep recurrent neural networks. IEEE.

[20] Aghajanian, et al. (2025). Longitudinal structural MRI-based deep learning and radiomics features for predicting Alzheimer's disease progression.

Alzheimer's Research & Therapy. 2025

[21] Loh, et al. (2021). Application of deep learning models for automated identification of Parkinson's disease: a review (2011–2021). Sensors.

[22]Jean-Louis and Seixas. (2024). The value of decentralized clinical trials: Inclusion, accessibility, and innovation. Science.

[23] Sinha, et al. (2024). Decentralisation in clinical trials and patient centricity: benefits and challenges. Pharmaceutical Medicine.

[24] Burger, et al. (2023). Opportunities and challenges with decentralized trials in Neuroscience. Biometrical Journal.

Q&A

Jianchang Lin (Takeda)