**The Prevalence and Impact of Phase 2 Trial Bypass in Neurology Drug Development**

Hannah Moyer, BSc1, Robyn Mellet, BSc1, Maya McKeown1, Karine Vigneault1, Jason Karlawish, MD2, Erika Augustine, MD3 Lon Schneider, MD4, Jonathan Kimmelman, PhD1

1. Department of Equity, Ethics and Policy, McGill University, Montreal, QC Canada

2. University of Pennsylvania

3. Kennedy Krieger Institute

4. University of Southern California

\* Corresponding author. Email: [jonathan.kimmelman@mcgill.ca](mailto:jonathan.kimmelman@mcgill.ca) Phone: (514) 953 3306; 2001 McGill College Ave, Montreal QC, H3A 1G1

**Word Count:** ~x

**Abstract**

**Introduction**

In neurology, P2 trials are primarily used to optimize dose and schedule1–6 and to map out the safety and tolerability of the treatment regimen under investigation.1,2,4,6. In addition, these trials are often designed to show the proof of concept behind the treatment, such as investigating whether it has the desired biological effect1,2,4 However, methods to provide proof of concept often rely on surrogate endpoints with little evidence that they are sensitive or reliable.7,8 Beyond proof of concept, showing signs of clinical efficacy in P2 trials is desirable, but often very difficult 1,2,6 For example, there are very few established clinical endpoints in early Alzheimer’s disease, partially due to the chronic nature of the disorders which prolongs the duration of clinical trials significantly compared to acute disorders.9 Due to the limitations associated with clinical endpoints, guidelines in ALS and AD research suggest that P3 trials can be initiated after receiving information on safety and tolerability, dose, proof of concept, all without clear clinical efficacy signals.1,9.

**Methods**

P3 Trial Sample

The ClinicalTrials.gov search to identify our sample of drugs was constructed using a list of terms for the following determined neurological diseases: Alzheimer's disease, Parkinson disease, Amyotrophic lateral sclerosis, Huntington's disease, Multiple sclerosis (RRMS and PMS), Migraine, Headache, Epilepsy, TBI and Stroke recurrence. All phase 3 trials with actual primary completion dates from January 1, 2011- January 1, 2021 using these terms were downloaded from ClinicalTrials.gov for screening.

Inclusion criteria was a) listed in ct.gov as “Phase 3, Phase 2/3, Phase 2b/3. Randomized, parallel group, or sequential, controlled trial; b) control is either placebo or another treatment (not a different dose of same drug; c) must test a drug or biologic; g) must be the first phase 3 trial for the treatment/indication pair registered on clinicaltrials.gov (unless there are phase 3 trials that are started within a year of each other and not completed): f) at least one US or CAD, EU, UK, Australian research site, g) Investigated a treatment for the included neurological conditions either treating the condition itself or a symptom of the condition that is widely used as measure of disease modification of the condition (as determined by consultation with neurologists). The exclusion criteria was a) head-to-head trials of standard of care interventions as the primary analysis, b) primary purpose is diagnostic or screening c) includes healthy volunteers, d) or Withdrawn (i.e. no patients enrolled).

Phase 3 trial publications were first searched for on ClinicalTrials.gov. For ClinicalTrials.gov records where no publication was linked, we conducted Google Scholar and OVID searches to find study publications for each trial. If multiple publications/abstracts were found, the primary publication will be chosen (i.e. the publication that reports full primary-endpoint results). All trials will need to have a publication of their results to be included. Publications only containing interim results will not be used unless the study was terminated at interim analysis.

Matching P3 Trials to Prior P2 Trials

To the introduction, conclusion, and research sections in the Phase 3 publications were searched for P2 trials. If none meet the matching criteria (see below), we searched TrialViewer (ClinicalTrials.gov) for additional P2 trials. If there are still no matches, we searched for P2 trials using google scholar, MEDLINE and EMBASE via OVID. For approved drugs, drugs@FDA documents will be used to check that we correctly matched P2 trials to P3 trials. As a last resort, corresponding authors or sponsors will be emailed to query about possible phase 2 trials.

To determine if a P2 trial was eligible to be a match, it had to have a primary start date that was one year earlier than primary start date of the phase 3 study in our sample, as indicated by ClinicalTrials.gov (or the recruitment start date of the publication if registration date was unavailable) and investigate the same drug or biologic in the same condition (see supplement for details).

Extractions

From each Phase 2 trial we extracted the positivity, the number of patients, and trial duration. To be deemed positive, P2 trials must have a primary clinical efficacy endpoint and be positive on that endpoint based on what was pre-specified in the trial. Alternatively, P2 trials were deemed to be ambiguous when a primary clinical endpoint was nonpositive or the trial used a biomarker/safety primary endpoint that is not validated.

From each Phase 3 trial we extracted termination status, positivity status, SMD on primary efficacy endpoint (if available), withdrawals due AEs, approval status, funding (industry vs nonindustry), number of patients, trial duration, and phase.

Prevalence of Bypassing

P3 trials were then put into the following groups: Preceded by P2: have a matched P2 trial that is positive per our definition above, ambiguous P2 bypass: Had a matched P2 trial that was nonpositive per our definition above, or true P2 bypass: Did not have a matched P2 trial.

To determine whether bypass was associated with any variables compared the proportions of trials that were not preceded by P2 on the following variables: Funded by industry vs. non-industry, condition was severe vs. non-severe condition (operationalized based on 5-year mortality or disability), and approval status.

When trying to determine whether a P2 trial could have provided preliminary efficacy evidence for the design of a P3 trial, we wanted to determine which surrogate endpoints we would consider “validated surrogates of efficacy.” These surrogates should and commonly used in phase 2 trials in that indication because of time constraints or other limiting factors. Makes sense mechanistically and has been validated in a P3 trial of a similar drug showing efficacy is associated with it. For example, we considered number of lesions as a validated surrogate endpoint in MS trials. However, we did not consider any AD surrogate endpoint to meet this criterion. The following endpoint is in our sample, and we wondered if you could tell us what you think. You can make notes in the table below:

Impact of Bypass on Risk and Benefit

Pos and termination is the most impt because those terminated don’t have results often

Restricting our analysis to those indication areas where there are at least 3 trials in the bypass and non-bypass group and within these indications, we performed a chi-squared test between the proportion of P3 trials in the bypass group vs non-bypass group and the P3: a) rate of termination, b) positivity on the primary endpoint

We performed a pooled meta-analysis with subgroup contrast between the bypass group vs non-bypass group for the following two variables. Pooled SMDs for efficacy outcomes, where trials involve continuous outcomes10 RR for withdrawal related adverse events for all trials in sample.

We found the average number of patients enrolled in phase 3 trials that are ultimately nonpositive on their primary endpoint or stopped early due to safety or futility that bypassed phase 2 evidence and compared this to the average number of patients in phase 2 trials in our sample. We will present the difference between these averages as the number of patients that could have potentially been spared if bypassing had not occurred (assuming the nonpositive result could have been discovered had a P2 trial been run). We will perform a similar analysis, estimating trial duration (e.g. patient years).

Ambigious is where we assume there is learning from a p2 trial but did not provide definitive efficacy evidence. This could be a proof of concept trial. This is called the “learn zone” of clinical research and likely does lead to support for P3 trial.11

Statistical Analyses

DO

Sensitivity analysis excluding p2/3

**Results**

Sample of Index P3 trials

After applying our inclusion and exclusion criteria, 91 trials were included in our sample. The vast majority of these trials were in Alzheimer’s disease (24 trials), and Migraine (22 trials).

Bypassed proportions as per both definitions (preceded vs ambig+bypass) and (preceded+ambig vs bypass) discussed here… **See Table 1**

Ambigious is imprt bc getting lots of information from trial outside of clinical positive result-still learning something

Risk and Benefit of P2 Bypassing

Using the (preceded vs ambig+bypass) definition, the positivity and termination rate for trials overall, in the preceded group, and the bypassed group are displayed for all indications in **Table 2**. Chi squared analyses will be conducted in indications where there were at least three trials in both the preceded and bypass group to determine if there was a difference in any of these groups.

Moral Economy

Phase 3 trials that are non positive or terminated for safety/futility avg number of patients and time duration of trial

**Discussion**

Limitations

P2/3s are put into preceded category although unsure if P2 had a clinical positive endpoint-different threshold to move to P3

Some phase2s found sig safety concerns and were terminated but we didn’t capture this

P2/3s letting them be evidence for themselves but don’t really know if they proceeded-especially ones without papers

**Declaration of Interest:** JK received consulting fees from Amylyx Inc. Authors declare no other potential conflicts of interest.

**Data Sharing:** Data will be available on Open Science Framework

**Acknowledgments:**

**Funding**: This work was funded by CIHR.

**Figure 1** -Prisma Flow Diagram for P3 Solid Tumor Trial Sample

Trial records identified from ClinicalTrials.gov (n = 1169)

**Identification**

**Records excluded semi-automatically (n=627):**

• Without an “actual” primary completion date (n=139)

• Non-randomized (n=215)

• Small sample size (n=82)

Trial Status is withdrawn (n=0)

• Primary purpose is diagnostic, screening, or basic science (n=3)

Not testing correct form of treatment (n=53)

• Healthy volunteers included (n=24)

• No US/CAD/EU/Australian enrollment site (n=108)

Duplicates (n=3)

Studies manually assessed for eligibility (n =542)

**Screening**

**Studies excluded manually (n=):**

• Intervention did not match our criteria (n=37)

Comparator did not match our criteria (n=13)

• Indication did not match our criteria (n=135)

• Not the first P3 trial in drug/indication pair (n=123)

First symptom check (

No results available (n=35)

• No primary efficacy endpoint (n=)

Treating a symptom (n=)

Phase 2/3 that did not continue to P3 portion (n=

Studies included in review (n =91)

**Included**

Table 1: Prevalence of Bypassing

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Indications | Overall  (N) | Type of supporting evidence | | |
| **Preceded by Positive P2**  **(N, %)** | **Preceded by Ambiguous P2**  **(N, %)** | **Bypassed P2**  **(N, %)** |
| All indications | 91 | 32, 35% | 31, 34% | 28, 31% |
| Alzheimer's disease | 24 | 7, 29% | 10, 42% | 7, 29% |
| Parkinson's disease | 10 | 3, 30% | 4, 40% | 3, 30% |
| Amyotrophic lateral sclerosis | 6 | 3, 50% | 3, 50% | 0, 0% |
| Huntington's disease | 5 | 0, 0% | 4, 80% | 1, 20% |
| Multiple sclerosis | 5 | 0, 0% | 1, 20% | 4, 80% |
| Migraine | 22 | 14, 64% | 5, 23% | 3, 14% |
| Headache | 3 | 0, 0% | 0, 0% | 3, 100% |
| Epilepsy | 7 | 1, 14% | 1, 14% | 5, 71% |
| TBI | 4 | 3, 75% | 0, 0% | 1, 25% |
| Stroke | 5 | 1, 20% | 3, 60% | 1, 20% |

Table 2: Positivity and Termination Rate and Bypass

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Indications | Positivity Rate of P3 | | | Termination Rate of P3 | | |
| **Overall (N=91)**  **(N, %)** | **Type of supporting evidence** | | **Overall (N=91)**  **(N, %)** | **Type of supporting evidence** | |
| **Preceded by P2**  **(N, %)** | **P2 bypass**  **(N, %)** | **Preceded by P2**  **(N, %)** | **P2 bypass**  **(N, %)** |
| All Indications | 38, 42% | 17, 53% | 21, 36% | 20, 22% | 6, 19% | 14, 24% |
| Alzheimer's disease | 1, 4% | 0, 0% | 1, 6% | 11, 46% | 3, 43% | 8, 47% |
| Parkinson's disease | 4, 40% | 1, 33% | 3, 43% | 2, 20% | 0, 0% | 2, 29% |
| Amyotrophic lateral sclerosis | 1, 17% | 1, 33% | 0, 0% | 1, 17% | 1, 33% | 0, 0% |
| Huntington's disease | 1, 20% | 0, 0% | 1, 20% | 2, 40% | 0, 0% | 2, 40% |
| Multiple sclerosis | 3, 60% | 0, 0% | 3, 60% | 0, 0% | 0, 0% | 0, 0% |
| Migraine | 21, 95% | 14, 100% | 7, 88% | 0, 0% | 0, 0% | 0, 0% |
| Headache | 1, 33% | 0, 0% | 1, 33% | 0, 0% | 0, 0% | 0, 0% |
| Epilepsy | 6, 86% | 1, 100% | 5, 83% | 1, 14% | 0, 0% | 1, 17% |
| TBI | 0, 0% | 0, 0% | 0, 0% | 2, 50% | 2, 67% | 0, 0% |
| Stroke | 0, 0% | 0, 0% | 0, 0% | 1, 20% | 0, 0% | 1, 25% |

Table 3. Candidate predictors of bypass

|  |  |  |  |
| --- | --- | --- | --- |
| Candidate Predictors | Overall (N=91)  (N, %) | Type of supporting evidence | |
| **Preceded by P2**  **N=32**  **(N, %)** | **P2 bypass**  **N=59** |
| **(N, %)** |
| Pharmaceutical sponsor | 68, 75% | 27, 84% | 41, 69% |
| Approved | 72, 79% | 23, 72% | 49, 83% |

Table 4. Moral economy number of patients and duration of running a P3 or P2 to find a non-positive result

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | All P3 | Nonpositive P3 | Terminated P3 | Phase 2 |
| Avg Number of Patients | 949 | 798 | 947 |  |
| Avg Trial Duration (days) | 1071 | 1077 | 1021 |  |

Add bypass here?

Figure 1 SMD 

Figure 2: WdAE