**The Prevalence and Impact of Bypassing Phase 2 Trials 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**

Neurological conditions are a major cause of death and disability globally.12 Despite being one of the most prevalent disease areas in modern life,2 neurological drug development has proven more challenging than others, with some indications lacking any established disease-modifying standard of care (SOC).3 Factors such as trouble ensuring that the drug is transported across the blood-brain barrier, reliance on endpoints that lack validation, and the long accumulative nature of the conditions together create an area of drug development where potential treatments have a low chance of success.4–63,78 However, positive results would have a massive impact on the experience of millions of patients,1 emphasizing the need for innovative modifications to the drug development process to get treatments to patients faster and to increase the incentives for companies to invest in their development.

Drug development for neurological conditions typically involves three sequential phases of clinical trials aimed at gaining FDA approval. Phase 1 (P1) trials are primarily focused on gathering pharmacological information. Phase 2 (P2) trials usually aim to understand safety and dose information while beginning to investigate efficacy. Together, P1 and P2 trials are called the “learn zone” of clinical research and support later trial designs.1 Phase 3 (P3) trials focus on confirming efficacy information in large numbers of patients on the treatment for longer periods of time.

One method used to potentially accelerate drug development is to initiate P3 trials without positive evidence from a P2 trial. We call this strategy “P2 bypass.”2 This practice has been discussed in the literature,3,4 although its prevalence is unknown as it has been described as both “rare”5 and the strategy of “many” sponsors.6 In cancer drug development, we found that nearly half of P3 trials investigating treatments for solid tumors bypassed P2.2

However, bypassing P2 trials and the information gained from them may diminish the chance of finding a positive result in a P3 trials because they are started with a lower amount of evidence on efficacy and safety. In addition, P3 trials that bypass P2 are not likely to have optimized dose and population details that make up the “intervention ensemble”, the package of variables surrounding the treatment that must be researched to make it clinically meaningful.55 Information from P2 trials can also guide “go/no-go” decisions for further testing to limit waste in drug development and weed out drugs that are not likely to be successful early in the development process.26,4912,51

Bypassing a P2 trial may be associated with diminished benefit and/or higher risk for patients participating in P3 trials. Our paper investigating P2 bypass in oncologic drug development found that patients in P3 trials that were not supported by P2 trials had significantly worse survival outcomes.65 Thus, clinical equipoise may be threatened for a P3 trial designed with little prior evidence.

In what follows, we created a sample of P3 trials registered on ClinicalTrials.gov and estimated the prevalence of P2 bypassing in ten neurological conditions. Secondly, we investigated whether bypassing positive clinical P2 evidence impacted P3 trial success. These results will help guide decision-making as to whether bypassing P2 trials is an efficient method of getting effective drugs to patients quickly.

**Methods**

P3 Trial Sample

We created our sample of P3 trials using a list of search terms on ClinicalTrials.gov for the following neurological diseases: Alzheimer's disease, Parkinson disease, Amyotrophic lateral sclerosis, Huntington's disease, Relapsing Multiple Sclerosis, Progressive Multiple Sclerosis, Headache, Epilepsy, TBI and Stroke. All P3 trials with actual primary completion dates from January 1, 2011- January 1, 2021 were downloaded from ClinicalTrials.gov for screening. Although we will refer to all trials in our sample as P3 trials, P2/3 trials were also included to provide an accurate estimate of the current clinical trial landscape.

Notable inclusion criteria included a) treatment involving a drug or biologic; b) at least one US or CAD, EU, UK, Australian research site, and d) disease modifying treatment or targeting a symptom of the condition that is widely used as measure of disease modification of the condition (as determined by consultation with neurologists). We excluded trials with the following features: a) the primary purpose was diagnostic or screening; b) included healthy volunteers, c) were preceded by a P3 or P4 trial that started more than a year before.

We first searched for P3 trial publications on ClinicalTrials.gov. When publications were not linked, we conducted searches on Google Scholar and OVID (MEDLINE and EMBASE). If we were unable to find any publication of results, we used results deposited on ClinicalTrials.gov.

Matching P3 Trials to Prior P2 Trials

For every P3 trial in our sample, we searched for “matched” P2 trials. For a P2 trial to be an eligible match for the P3 trials in our sample, it had to 1) have started one year earlier than the P3 trial, as indicated by primary start dates on ClinicalTrials.gov and 2) investigate the same treatment in the same condition. To find potential matches, we searched the introduction, conclusion, and research sections in P3 publications. If no cited trials met the matching criteria (or no P3 publication was available), we searched ClinicalTrials.gov for additional P2 trials. For approved drugs, we used FDA approval documents to check that we correctly matched P2 trials to P3 trials. When we could not find any matched P2 trials, corresponding authors or sponsors were queried about possible P2 trials.

Extractions

We extracted positivity status from each matched P2 trial. To be deemed “positive”, P2 trials must have had a primary clinical or “validated surrogate” efficacy endpoint and be positive on that endpoint based on what was specified in the trial. For a surrogate endpoint to be considered a “validated surrogate of efficacy,” it had to be commonly used in P2 trials in that indication because of time constraints or been validated in a P3 trial of a similar drug. For example, we considered number of gadolinium-enhancing lesions to be a validated surrogate endpoint in relapsing multiple sclerosis trials while we did not consider any Alzheimer’s disease surrogate endpoints to meet this criterion. This was determined in collaboration with neurologists.

From each P3 trial in our sample, we extracted termination status, positivity status, SMD on primary efficacy endpoint, proportion of patients who withdrew due adverse events in each arm, approval status, funding (industry vs non-industry), and phase.

Analyses

Prevalence of Bypassing

Our primary outcome was to estimate the prevalence of bypassing across neurological indications in our sample. To do so, we found the proportion of P3 trials that were launched using four different trajectories. The first trajectory was: “Preceded by a positive P2:” P3 trials that were preceded by a P2 trial that was positive on a primary clinical or validated surrogate endpoint. The second trajectory was “Preceded by an Ambiguous P2:” P3 trials that were preceded by a P2 trial that provided evidence other than that from primary efficacy result. This trajectory was split into two: a) “Non-positive:” P3 trials were preceded by a P2 trial that was non-positive on clinical or validated surrogate endpoints; b) “Not focused on efficacy:” P3 trials were preceded by a P2 trial that investigated proof of concept endpoints, only investigated safety, or used non-validated surrogate endpoints. The final trajectory was “True bypass”: P3 trials were not preceded by a P2 trial in the same indication with the same drug. We considered the first trajectory “Non-bypass” and the remaining trajectories were considered “P2 Bypass”.

As a secondary analysis, we investigated whether P2 bypass was associated with industry funding or the approval status of the experimental treatment at the time of trial initiation using a fisher-exact test.

Impact of Bypass on P3 Trial Success and Patient Risk/Benefit

For a secondary analysis, we investigated whether bypassing positive clinical evidence from P2 trials affected P3 trial success. To do so, we performed a fisher-exact test to determine whether bypassing impacted the rate of positivity on P3 primary outcomes or P3 termination due safety or futility. Additional secondary analyses included a pooled meta-analysis of standardized mean differences (SMD) and risk ratios (RR) for withdrawal-related adverse events with subgroup contrasts between the P3 trials that bypassed vs those that did not bypass.

Statistical Analysis

Fisher-exact tests were performed using the “prop.test” R function.1 Significance was determined using nominal significance level of 0.05 for all analyses. To compare SMD in efficacy outcomes between trials that bypassed and those that did not, we used the function “metagen” from the “meta” R package.2 This analysis was restricted to indication areas where there are at least 3 trials in the bypass and non-bypass group that had SMD results using the same scale. To compare whether risk of withdrawal due to adverse events was impacted by bypassing, we used the function “metabin” from the “metafor” R package.3 We used the two-tailed *p-*value of Cochran's Q for subgroup difference to investigate significance.

See supplement for more methodological details and protocol deviations. All extractions were performed in duplicate, and consensus was sought from JK. Our protocol was registered at <https://osf.io/crf62/>

**Results**

Sample of Index P3 trials

After applying our inclusion and exclusion criteria, 113 trials were included in our sample (**see Figure 1**). The majority of these trials were investigating treatments for Alzheimer’s disease (27%), and headache (23%). Most trials were funded by industry (83%) and were investigating treatments that were not approved at the time of trial initiation (81%) (**See Table 1).**

Prevalence of Bypassing

Overall, bypassing positive efficacy evidence from a P2 trial was common in neurologic drug development (48%). This practice was most common for trials investigating treatments for stroke (83%), progressive multiple sclerosis (75%), and Huntington’s disease (75%) and was least common in relapsing multiple sclerosis (6%) and headache (27%) trials (**see** **Table 2).** The prevalence of P2 bypass was not associated with industry funding or approval status (p= 0.21, p=0.47 respectively) (**see Table 3**).

Bypassing and P3 Trial Success

In our sample, P3 trials that bypassed P2 were significantly less likely to be positive on their primary outcome than trials that were preceded by positive efficacy evidence from a P2 (Non-bypass: 57% vs Bypass: 32%, p=0.01). In addition, the rate of P3 trial termination due to safety or futility was non-significantly higher in the group that bypassed P2 (Non-bypass: 15% vs Bypass: 28%, p=0.11) **(see** **Table 4).**

Patient Risk and Benefit of P2 Bypassing

In our sample, the only indication with a sufficient number of trials in each subgroup that reported the same outcome was Alzheimer’s disease. The pooled SMDs for The Alzheimer’s Disease Assessment Scale–Cognitive Subscale (ADAS-Cog) were not significantly different between trials that bypassed and those that did not (p=0.13). However, there was a trend toward better outcomes in trials that bypassed P2 (**see** **Figure 2)**. Similarly, pooled RRs for withdrawals due to adverse events were not significantly different between trials that bypassed and those that did not (p=0.31). There was a clear trend toward higher risk of withdrawals due to adverse events in P3 trials that bypassed P2 in our full sample and in trials investigating treatments for Alzheimer’s disease. The opposite trend was shown in trials investigating treatments for headache (**see** **Figure 3).**

**Discussion**

**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.

**Figures and Tables**

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

Trial records identified from ClinicalTrials.gov (n = 1188)

**Identification**

**Records excluded semi-automatically (n=647):**

• Without an “actual” primary completion date (n=142)

• Non-randomized (n=216)

• Small sample size (n=83)

Trial Status is withdrawn (n=1)

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

• Incorrect intervention or indication (n=80)

• No US/CAD/EU/Australian enrollment site (n=119)

• Duplicates (n=3)

**Screening**

Studies manually assessed for eligibility (n =541)

**Studies excluded manually (n=428):**

• Intervention did not match our criteria (n=51)

• Comparator did not match our criteria (n=17)

• Indication did not match our criteria (n=146)

• No primary efficacy endpoint (n=49)

• Not the first P3 trial in drug/indication pair (n=159)

• Phase 2/3 that did not continue to P3 portion (n=6)

Studies included in review (n=113)

**Included**

**Table 1. Characteristics of the Sample**

|  |  |
| --- | --- |
| Indications | Number of P3 trials  N=113 (%) |
|
|
| All indications | 113 |
| Alzheimer's disease | 30 (27) |
| Parkinson's disease | 10 (13) |
| Amyotrophic lateral sclerosis | 5 (4) |
| Huntington's disease | 4 (4) |
| Relapsing Multiple sclerosis | 16 (14) |
| Progressive Multiple sclerosis | 4 (4) |
| Headache | 26 (23) |
| Epilepsy | 7 (4) |
| TBI | 5 (4) |
| Stroke | 6 (5) |
| General |  |
| Pharmaceutical funder | 94 (83) |
| Pre-approval status | 92 (81) |
| Positive primary endpoint | 49 (45)\* |
| Terminated for safety or futility | 24 (21) |
| Phase 3 | 100 (88) |
| Median sample size (IQR) | 835 (706) |

\*Out of 108 trials with primary results available

**Table 2: Prevalence of Bypassing**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Indications | Overall  (N) | Non-Bypass | Bypass | | |
| **Preceded by Positive P2**  **(N, %)** | **Preceded by Ambiguous P2**  **(N, %)** | | **True Bypass**  **(N, %)** |
| **Nonpositive** | **Not focused on efficacy** |
| All indications | 113 | 59 (52) | 19 (17) | 15 (13) | 21 (19) |
| Alzheimer's disease | 30 | 9 (30) | 8 (27) | 7 (23) | 6 (20) |
| Parkinson's disease | 10 | 5 (50) | 0 (0) | 4 (40) | 1 (10) |
| Amyotrophic lateral sclerosis | 5 | 3 (60) | 2 (40) | 0 (0) | 0 (0) |
| Huntington's disease | 4 | 1 (25) | 2 (50) | 1 (25) | 0 (0) |
| Relapsing multiple sclerosis | 16 | 15 (94) | 0 (0) | 1 (6) | 0 (0) |
| Progressive multiple sclerosis | 4 | 1 (25) | 1 (25) | 1 (25) | 1 (25) |
| Headache | 26 | 19 (73) | 4 (15) | 0 (0) | 3 (12) |
| Epilepsy | 7 | 2 (29) | 1 (14) | 0 (0) | 4 (57) |
| TBI | 5 | 3 (60) | 0 (0) | 1 (20) | 1 (20) |
| Stroke | 6 | 1 (17) | 1 (17) | 0 (0) | 4 (67) |

**Table 3. Candidate Predictors of P2 Bypass**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Candidate Predictors | Overall (N=113)  (N, %) | Type of supporting evidence | | P-values  non-bypass vs bypass\*\* |
| **Non-Bypass**  **N=59**  **(N, %)** | **Bypass**  **N=54**  **(N, %)** |
| Pharmaceutical funder | 94 (83) | 52 (88) | 42 (78) | 0.21 |
| Approved | 21 (19) | 9 (15) | 12 (22) | 0.47 |

**Table 4: Overall Positivity and Termination Rate and Bypass**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Non-Bypass | Bypass | | | P-values  non-bypass vs bypass\*\* |
| **Preceded by Positive P2**  **N (%)** | **Preceded by Ambiguous P2**  **N (%)** | | **True Bypass**  **N (%)** |
| **Nonpositive** | **Not focused on efficacy** |
| Positivity Rate\* | 33/58 (57) | 5/19 (26) | 3/14 (21) | 8/17 (47) | .01 |
| Termination Rate | 9/59 (15) | 6/19 (32) | 3/15 (20) | 6/20 (30) | .11 |

\*Trials were only included in the positivity analysis if they had primary results available

\*\*fisher exact test between trials in non-bypassed trajectories vs bypassed trajectories:

**Figure 2: SMDs for ADAS-cog Pooled Subgroup Analysis.**

****

**Chart, box and whisker chart

Description automatically generated**

**Figure 3: RRs for WdAe Pooled Subgroup Analyses**

