Supplement

**P3 Sample Creation**

CT.gov SEARCH PARAMETERS For Phase 3 Trials:

1. Condition or disease: Alzheimer disease OR Alzheimer's disease OR Alzheimer Dementias OR Dementia of the Alzheimer's type OR dementia alzheimers OR Dementia of Alzheimers Type OR dats OR Alzheimer Type Dementia OR Senile Dementia OR Alzheimer Syndrome OR AD OR Parkinson disease OR Parkinson's disease OR PD OR Parkinson OR Primary Parkinsonism OR Paralysis Agitans OR Shaking palsy OR ALS OR Amyotrophic lateral sclerosis OR Gehrig Disease OR Motor neurone disease OR Charcot disease OR Huntington disease OR Huntington's disease OR Huntington's chorea OR Chronic progressive hereditary chorea OR MS OR Multiple Sclerosis OR MS (Multiple Sclerosis) OR Disseminated sclerosis OR Migraine OR Cephalalgia OR Head pain OR Pain in head OR Cephalgia OR Headache OR Epilepsy OR epileptics OR seizure disorders OR epilepsia OR TBI OR Traumatic Brain Injury OR brain traumas OR Traumatic encephalopathy OR brain injuries traumatic OR traumatic brain damage OR Brain damage OR cerebral damage OR injury brain OR cerebral injury OR Stroke OR Cerebrovascular accident OR cerebral vascular accident OR Apoplexy OR Brain attack OR Brain Vascular Accident OR TIA (Transient Ischemic Attack) OR Transient Ischemic Attack OR intracerebral haemorrhage OR subarachnoid haemorrhage
2. Study type: “Interventional Studies (Clinical Trials)”
3. Status of recruitment: no restriction (looking for Actual primary completion dates, so likely mostly Completed/Terminated/Active not recruiting but completed- checked filtered results to see)
4. Phase: 3
5. Study start date: no restriction
6. Primary completion date: 01/01/2011-01/01/2021

The end range was chosen to allow one year between primary completion and depositing results as per the Final Rule.1 Our objective was to have at least 100 phase 3 trials but we saturated the sample for a full decade of Phase 3 trials. The target minimal sample size of 100 is selected because, for a primarily descriptive study, it seems likely to deliver a reasonably robust estimate of the prevalence of phase 3 bypass. Assuming 30% trials involve phase 2 bypass, availability of 30 trials involving bypass provides a reasonable starting point for secondary objectives for a first ever exploration of the prevalence of bypass.

SEMI-AUTOMATIC SCREENING (using excel filters) For Phase 3 Trials:

1. Primary completion date: checked that type is “Actual” and not “Anticipated”
   1. Excluded, \*unless\* trial had an “Actual” overall completion date;
2. Trial design: excluded if trial was labelled as:
   1. “Non-randomized” in randomization field;
   2. “Single group assignment” in “Model” field;
   3. 1 in “Arms” field;
3. Trial size: <30
4. Trial status: exclude if the trial recruitment status was:
   1. Withdrawn (i.e. no patients enrolled);
5. Indication: excluded if primary purpose is
   1. Diagnostic;
   2. Screening;
   3. Basic Science
6. Intervention/Indication: excluded if trial:
   1. Did not include at least one intervention that was classified as a “Drug” or “Biological” “ Dietary supplement” or “genetic” (“Other” and “combination product” is manually checked); ie exclude procedure or behavioral or device or radiation
   2. Included healthy volunteers;
7. Trial Location: exclude if the trial does not have a
   1. US or CAD UK, EU, Australian

MANUAL SCREENING For Phase 3 Trials:

1. Intervention: Exclude if the intervention is
   1. surgery/behavioral/device/conditioning of stem cells/procedure
   2. extension, discontinuation studies, phase 1/2/3
   3. head to head (trials pitting two approved SOC interventions against each other) or if there are more than two options for the experimental arm (ake “any anticoagulant”)
   4. treating a second condition in our conditions (ie infection in PD patients) (immune responses to vaccines)
   5. biosimilar against what it is biosimilar to
2. Comparator
   1. must use a comparator that is either placebo or another treatment (as opposed to another dose of same drug (no historical controls)
3. Indication-Must investigate treatment for the below conditions exclusively:
   1. Alzheimer's disease
      * + 1. Excluded trjals investigating treatments for:

Healthy people with AD mutations

or MCI without pathologic characteristics of AD

* + - * 1. Included trials investigating treatment for:

Trials investigating MCI with pathologic characteristics of AD (prodromal) were included

Mild-severe AD (however defined)

* 1. Parkinson disease,
  2. Amyotrophic lateral sclerosis,
  3. Huntington's disease,
  4. Relapsing Multiple sclerosis, Progressive Multiple sclerosis,
     + - 1. CIS was not included
  5. Headaches,
  6. Epilepsy,
     + - 1. Antiseizure considered the same as antiepileptic
  7. TBI and
  8. Stroke
     + - 1. Must be in patients who have had a stoke looking at recurrence or recovery

1. No earlier Phase 3 trial: Must not be preceded by a phase 3 or 4 trial that had more than a year of progress. We used TrialViewer to search for all earlier phase 3 trials of our experimental drug-of-interest. We used the following rules for determining if earlier phase 3 trials counted as evidence for the trial in our sample.
   1. If the intervention is a change in the administration of a different drug, P3 trials investigating the other drug are not counted as prior evidence
   2. If the intervention is treating a symptom of a condition
      1. If there are prior trials investigating the same treatment in the same condition treating the same symptom, or has outcomes looking at the general condition,-these can be used as prior evidence
      2. If there are prior trials investigating a different symptom in the same condition and does not look at general condition modification or the symptom in our trial-this is not prior evidence.
   3. Prior trials that investigate treatments in preclinical populations are not used as prior evidence
      1. Exp: CIS, people with AD mutation
   4. To count as a prior P3 evidence for trials in our sample, earlier trials
      1. did not need to be exclusively in that indication
      2. could be Phase 4 or 3
      3. the same intervention could be in control or exp arm, just needed to have been studied in a phase 3 trial in that indication before
      4. Trials investigating could treatment in children could be evidence for adults and other way around
      5. Same Day year before is earlier evidence
   5. Adjunctive and monotherapy trials could be matched with each other. Trials investigating combination therapies only matched to combination trials.
   6. RRMS and Progressive MS were treated separately, they could not be used as prior evidence for the other. If the trial was only SPMS, it could cite either RRMS or PMS.
   7. We did not check for the status of the previous trial.
   8. Variations of old drugs in the same indication were excluded
      * + 1. drugs that are SLIGHT variations on old drugs used in the same indication-either because they have made a small molecular change or changed the delivery mechanism because these trials have a heightened level of evidence available from trials on the original drug. This is considered earlier P3 evidence (when approved or either P3 trial in the same indication.
          2. criteria is reliant on citations-cant check everything same family of drug does not count
2. Trials were included if they had a primary endpoint that was a clinical efficacy endpoint widely used measures of disease modification of each disease area for phase 3 trials
   * + 1. Trials were excluded if they only had primary safety, tolerability or surrogate primary endpoints or primary endpoints looking only at a symptom that is not used as a measure of disease modification.
       2. Neurologist collaborators were queried: “Would you consider any of the following “widely used measures of disease modification of X” for phase 3 trials?”
3. Phase 2/3 trials that did not progress to the P3 portion were excluded.
   * + 1. We used how it is defined on ct.gov in the phase category unless they call themselves something else in the publication or the ct.gov record.

**Phase 3 results**

The Google Scholar search used NCT ID, Title (top-line & official), varying combinations of drug names, indication, and sponsor & investigator last name. The OVID search using MEDLINE and EMBASE used a combination of the search terms: drug names from the experimental arm (any synonym of the drug mentioned in ClinicalTrials.gov should be included), and indication as listed in ClinicalTrials.gov , and “Clinical trial”, and “Phase 3”.

Primary results are defined as reporting the results on at least one primary outcome with a significance test

performed. If this did not turn up primary publications of results, we used primary Ct.gov results. If there are no primary ct.gov results, we will use abstracts reporting primary results. We only used abstracts reporting interim results if the trial was terminated. Trials without results are included in the prevalence results but not in the positivity analysis (unless they were terminated at DSMB review-which would result in nonpositive classification).

**Matching P2 trials to P3 trials**

Matches were found from citations in P3 trials, clinicaltrials.gov, FDA approval documents, author solicitation, and Alzforum. P2/3 are put into the proceeded category automatically. If the P2 trial was terminated with no data or an extension study, it is not counted as prior evidence. Otherwise, to determine if a P2 trial was eligible to be a match, it had to have a primary start date that is a year or more 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). If the date that the P2 trial started is unclear, publication within/before the year that the P3 trial started is accepted. Expanded access trials, extension studies, non-prospective trials, and trials without any accessible results were not considered. In addition, P2 trials had to also match on:

1. Indication
   1. Phase 2 trials were only considered a match to the phase 3 trial in our sample if it is in the same condition. To ensure our approach for matching phase 2 and 3 trials was standardized and reproducible, we allowed any P2 trial in the same BROAD disease area count as an earlier phase 2 trial for the phase 3 trials in our sample. Our broad disease areas are Alzheimer’s disease, Parkinson’s disease, ALS, Huntington’s disease, Multiple sclerosis\*, Headache, Epilepsy, TBI and Stroke recurrence.

\*Relapse Remitting MS and Progressive MS were treated separately, and they could not cite the other as prior evidence. If the trial was only SPMS, it could cite either RRMS or PMS.

1. Intervention
   1. P2 trials had to investigate the same drug or biologic
      1. A trial that investigates a drug/biologic as a monotherapy cannot be used as prior evidence for a trial that is investigating the same drug in combination therapy (and vice-versa). Monotherapy evidence could be used for adjuvant P3s because it may just be a result of shifting populations from early line to late line patients. Adjuvant evidence could also be used for monotherapy P3 trials in our sample. Two adjuvant trials with different background drugs were accepted as potential matches.
         1. Adjuvant definition: Adjuvant trials were identified by the terms “adjuvant” or “add on”: new (Experimental) + old
         2. Combo definition: new + new or old + old
      2. slight variations on drugs-either because they have made a small molecular change or changed the delivery mechanism are considered as matches (unless of course, these old versions preceded to P3 trials in which case the trial with the new variation on the old treatment would be excluded from our sample.

Phase determination:

It was occasionally unclear as to which phase an earlier trial was. We used how it is defined on ct.gov in the phase category unless they call themselves something else in the publication or the ct.gov record. If still not defined, we used the following rules.

* P1-Not controlled and no efficacy endpoint
* P2-- Call itself dose-ranging, or proof of concept, All that does not fall into the other two categories (Can be controlled or not controlled, can have an efficacy endpoint, and has a small\* number of patients)
* P3- Controlled and have primary efficacy endpoints and large\* number of patients

\*Dependent on indication: for now: 300

Positivity of P2 matches

We used the definition of positivity provided by the trial including using sequential testing procedures regardless of whether they modified their primaries. Trials that were stopped but were then positive were considered positive. In rare cases, there was no distinct primary endpoint, the coders tried to determine the objective of the study to the best of their abilities. If there was no statistical analysis or definition of positivity, the trial was considered to be an ambiguous prior P2 that is not aimed at efficacy. When there were two primary analyses where one was positive and the other was not (inconsistent results), we used the following rules:

1. Co-primaries: When they stated that all primaries had to be positive for the trial to be positive, we called inconsistent results nonpositive. When they don’t change the sig level
2. Multiple primaries: In this case, each primary endpoint is tested at a significant level determined by the method for multiplicity adjustment or simply by the partition of the alpha levels. we called inconsistent results positive\* if; they call it coprimaries but adjust the primary as is common with multiple primaries we will treat as multiple
3. In cases where there were 2 dose groups are both considered primary analysis groups, we called inconsistent results positive, therefore not needing multiplicity adjustments for multiple dose arms.

Classification

Each P3 trial was classified into one of the following groups based on its prior evidence:

* + To be put into the preceded group
    - P2 trial that had a clinical primary endpoint that was positive as defined in the trial. Futility trials were included in this group
    - P2 trial that had an approved primary surrogate endpoint that was positive as defined in the trial. Surrogate endpoint will only be included in this group if
      * they are reasonably validated “biomarkers of disease pathophysiology”2 type 2 biomarkers.3 No AD surrogates were accepted. 2
        + They are used to evaluate efficacy, they are commonly used as a primary endpoint in phase 2 trials in that indication because of time constraints, and make sense mechanistically, and have been validated in a P3 trial of a similar drug showing clinical efficacy is associated with it
        + Allowed
        + The only surrogate that we considered reasonably validated was number of gadolinium-enhancing lesions for multiple sclerosis trials.

MS-Number of lesions

MS-responder rate, defined as the proportion of patients with ⩾95% peripheral CD19+ B-cell depletion from baseline within 2weeks

* + To be put into the ambiguous group: every other trial with a matched P2 trial
    - Non-positive: Had a P2 trial that was nonpositive on their primary clinical efficacy endpoint,
    - Not aimed at providing efficacy data: had a P2 trial that had a primary endpoint investigating surrogate endpoints (not approved), or investigating safety/tolerability etc, or without an analysis run to evaluate significance, definition of positivity, or tested inferiority. When it is unclear if the primary was efficacy-trials were put in this group.
  + To be put in the bypass group
    - No matched P2 trial. When we cant find pubs or results, these trials are put into the true bypass group.
    - These were confirmed with emails to authors when emails were available.
* If any p3 trial had more than one prior trial, the one closest to preceded will take priority. If there were trials with conflicting results (pos and negative), the p3 trial was considered to have positive evidence. This is because the presence of a trial with nonpositive results does not negate the presence of positive results and we do not want to disincentivize publishing non positive results.

**Extraction**

* Termination status
  + And whether it was due to futility or safety concerns. This was found on registration records or in publications
  + There were two that were terminated at interim and guessed it was futility—why we cant separate futility and safety for termination
* Positivity status
  + All P3 trials in our sample had to have a primary clinical endpoint. They were deemed to be positive if they were positive on a primary outcome Trials that were stopped but were then positive, were considered positive. If the trial was stopped by DSMB but no results were available, trials were deemed to be positive (although there was one case where this wasn’t the case).
    - We looked at p-values and the definition of positivity in the statistical analysis section to determine trial positivity. The same positivity results as above were used
* When available, we extracted SMD of efficacy endpoints and the withdrawals due to AEs
  + When there were different dose groups as the primary
    - Took the one that is first for hierarchical testing. If there truly was not one higher priority we took the higher dose. If one was added as an amendment, the original was taken.
  + SMDs-
    - Took the most adjusted p-value available. If
    - Headache
      * Endpoint we want
        + Overall Mean Change From Baseline in Number of Monthly headaches
        + Extract if they have weekly or monthly-priority on monthly
    - AD
      * Endpoint we want
        + ADAS-COG (13 if it was available). If lsm was clearly placebo-exp the neg sign was switched. We wanted exp-placebo
    - PD
      * Endpoint we want
        + UPDRS total scores (from subscales I-III or I-IV) extracted if available. Otherwise, combined UPDRS II+III or UPDRS III scores were extracted.
  + WdueAE
    - From ct.gov or consort documents in the publications. Where there was disagreement, the publication took priority. We tried to make this the same group as was taken for SMD but if it was not possible could be different (broader groups)
    - Took from baseline randomization as denominator (earliest that was available)
* Approval status
  + Trials were classified as pre/post approval for the treatment under investigation at the time of trial initiation (primary start date in registration)
    - Pre-approval = drug was approved after the primary start date or never approved  
      post-approval = drug was approved before the primary start date
  + Approval did not need to be in the same indication or delivery mechanism
  + In combination, if both are approved separately we call post approval
  + If the trial is looking at a new formulation for an old drug- the first formulation will be used for approval date
* Funding (industry vs nonindustry).
  + If no funder was available, we took the sponsor
  + When funded by gov but drug was supplied by pharma we called in nonpharma
* Did it move on to from P2 to P3 in P2/3?
  + Unless they specified that they did not or called it a P2, we will include

**Analyses**

Prevelance

All included trials

Bypass definition

All following analyses used the subgroup definitions of preceded vs (ambiguous+bypass)

Positivity/Termination

Positivity- this was performed for only trials that had results available to determine positivity

Termination-this was performed for all included trials

Pvalues were only used for all indications analysis because of the lack of power

Performed a Fisher Exact probability test to find pvalues

Secondary and bypass

Only done in approval, and funder. This was performed for all included trials

Performed a Fisher Exact probability test to find pvalues

SMD

Only performed when the primary endpoint was the same outcome measure in 3 trials in each subgroup. Was only AD and PD

Only taken when the mean difference was presented along with pvalues or Cis

WdAE

Only performed where data was available in 3 trials in each subgroup

**Stats**

Fisher-exact tests were performed using the “fisher.test” R function.17 wdAE we used the function “metabin” from the “metafor” R package.19

**Protocol Deviations**

* Did not search pubmed
* Positivity of p2 was not include times when there are multiple trials with conflicting results,
* Did not look at these variable
  + Phase 2/3 vs P2 (these were all preceded)
  + Pediatric vs Adult vs Mixed (all were adult)
  + Orphan disease (all were not orphan (except maybe HD)
  + Symptom (most were excluded)
  + Severity-too difficult to operationalize
* Did not do d (separated into true bypass and ambiguous bypass
* In addition, we will estimate the proportion of patients in our sample that were in P3 trials that bypassed.
* Moral econ-did not use neg used nonpositive
  + Did all trials not Within the five disease areas with the largest number of trials in our sample,
  + Did not do In addition, we will estimate the proportion of patients in our sample that were in P3 trials that bypassed.
* SMD and Wdae analyses needed 3 in each group with SMD numbers available
* SMD
  + The only surrogate that we considered reasonably validated was number of gadolinium-enhancing lesions for multiple sclerosis trials.
  + Put in surrogate
* Did not search ovid for the matches

Results

**Chart, box and whisker chart

Description automatically generated**