Title
Late-Stage Failures of Monoclonal Antibody Drugs: A Retrospective Case Study Analysis

Journal
Pharmacology, 105(3-4)

ISSN
0031-7012

Authors
Sun, Amy
Benet, Leslie Z

Publication Date
2020

DOI
10.1159/000505379

Peer reviewed

Late Stage Failures of Monoclonal Antibody Drugs:

A Retrospective Case Study Analysis

Amy Sun[a*] and Leslie Z. Benet[b]

[a]School of Pharmacy, University of California, San Francisco, CA, USA

[b]Department of Bioengineering and Therapeutic Sciences, Schools of Pharmacy and Medicine, University of California, San Francisco, CA, USA

[*]Present address: Product Development Regulatory - Program Management Genentech Inc., South San Francisco, CA, USA

Running Head: Late Stage Failures of Monoclonal Antibody Drugs:
A Retrospective Analysis

Corresponding Author: Leslie Z. Benet, 533 Parnassus Ave Rm U-68, UCSF Box 0912, San Francisco, CA 94143; Phone: 1 415 476 3853; E-Mail: leslie.benet@ucsf.edu

Keywords

late stage failure; lampalizumab; bococizumab; tabalumab; solanezumab; bavituximab; sirukumab;

Abstract

Aim: To analyze the late stage failures of monoclonal antibody drugs. The later a drug fails in development, the more time and expense is incurred by the sponsor. **Methods:** We review the late stage, Phase III, failures of 21 monoclonal antibody drugs between 2014 and 2019 using published and publicly available information to characterize the reasons for these failures. **Results:** In some cases the failures are unavoidable due to the lack of adequate science, but in others we characterize the causes of such failures and recommend how such failures may have been avoided. **Conclusion:** By learning from previous mistakes and adhering to the principles and recommendations provided, it is possible to avoid these common pitfalls, increasing the likelihood of success in phase III clinical trials and thus securing regulatory approval.

Words: 10,344

Tables: 2

Introduction

New drug development and approval is a highly time and resource intensive undertaking. Currently, it takes on average a decade or more to bring a drug to market and costs an estimated $2.6 billion (including the cost of failed drug candidates) [1]. The later a drug candidate fails, the more time and resources have been sunk into its development. For both large and small

companies, the consequences of late stage failure can be devastating [2]. The "fail early, fail fast" (or "fail early, fail cheap") strategy of many pharmaceutical companies seeks to identify and terminate poor drug candidates early in their development, in particular, prior to expensive, years-long phase III clinical trials [3, 4].

Nevertheless, over half of all drug candidates will ultimately fail in phase III trials, which are meant to be confirmatory trials [5]. The persistent phenomenon of late stage failures has provoked thoughtful investigation and rigorous analyses to uncover causes, trends, and learnings that can be carried forward to future clinical development programs [6]. In many instances, however, late stage failures are only broadly categorized into reasons of efficacy, safety, or commercial factors, without substantial explanation or detail [5, 7]. In other instances, information disclosures such as press releases may provide more specific reasons for clinical failure of a drug candidate (*e.g.*, failure to meet a primary endpoint of improving overall survival [8], increased rate of death and fatal infections [9], etc.), but do not address more than one therapeutic compound at a time or elucidate trends across compounds.

Therefore, a better and more detailed understanding is needed to guide future drug development efforts. This is especially true for biologics, which are more complex and expensive to develop than their small molecule counterparts, and which claim an increasingly larger role in healthcare delivery and costs [10, 11]. In 2016, biologics accounted for 25% ($232 million) of the global pharmaceutical market [12].

Accordingly, we sought to identify and evaluate in detail the reasons for late stage failure of large molecule, mAbs, drugs, and to determine whether there are distinct or underappreciated reasons for their failures in phase III. We provide herein specific evidence-based recommendations to steer biologics development away from common pitfalls, increasing the likelihood of success in phase III and ultimately securing regulatory approval.

Reasons for Failure

In Table 1 we list 21 large molecule mAbs drugs where publicly available information identifies these therapeutic agents as failing during/after phase III clinical trials, and where regulatory approval was not obtained.

[Table 1 should be here.]

Drug candidates can fail late in clinical development for a multitude of reasons. Here for 12 of these candidates we provide a more detailed analysis within the following categories: **1) Non-drug related strategic and commercial factors.** A common, and very unfortunate, reason for drug failure of both large and small molecular entities may be ascribed to strategic and commercial factors that drive decision-making in clinical development that are independent of the science. Here drug molecules are moved into phase III when they should have been stopped earlier in the drug development process. Sponsors ignore the "fail early-fail cheaply" dictum since the sponsor has no other candidates to move forward. We believe that bavituximab falls into this category. **2) Trial design that is not sufficiently rigorous in identifying a positive**

outcome. Lampalizumab is reviewed in this category. **3) Posthoc subgroup analysis and false biomarker identification.** We review the development of lampalizumab, onartuzumab and rilotumumab in this category. **4) Incomplete understanding of the disease pathway.** This is a very common situation in early drug development, especially for first-in-class moieties. This reason for failure applies to lampalizumab and all of the drugs to treat Alzheimer's disease; here solanezumab is reviewed. **5) Target engagement may not result in disease modification.** This is actually a subcategory of topic 4. Here we specifically review this reason for failure for lebrikizumab and tralokinumab, in addition to solanezumab. **6) Suboptimal dosing.** Bococizumab and tabalumab are discussed. **7) Proceeding directly from phase I to phase III.** This frequently employed, often unwise, strategy leads to many of the other categories enumerated here. But, we specifically review this for tabalumab. **8) Multiple simultaneous clinical trials.** This reason for failure does not allow lessons learned from one clinical trial to be applied to the next. Here bococizumab and tralokinumab are reviewed. **9) Improper inclusion criteria.** We review solanezumab. **10) Inappropriate escape criteria confound outcome measurements.** This can be viewed as a subcategory of category 2. We review the sirukumab and tremelimumab phase III studies. **11) A poorly run phase II study can lead to an unrealistic phase III design.** Bavituximab is reviewed. **12) Antidrug antibodies.** A very difficult problem to analyze and predict the relevance; here we review bococizumab. **13) Preclinical results that do not translate to clinical outcomes.** Another problem that can be difficult to predict, here we review solanezumab.

We first review lampalizumab, which failed in phase III clinical trials due to insufficiently stringent phase II trial design, incomplete understanding of disease pathway, and retrospective

subgroup analysis from phase II results that led to false biomarker identification. Onartuzumab and rilotumumab are also presented as points of comparison to lampalizumab as far as identification of false biomarkers in phase II that resulted in inappropriate study populations in phase III.

Lampalizumab

Geographic atrophy (GA) is characterized by localized atrophy of the retina occurring in late stages of dry age-related macular degeneration (AMD). There is currently no treatment that can slow, stop, or reverse the course of the disease.

Lampalizumab is a monoclonal antibody fragment developed by Genentech and tested in clinical trials for treatment of GA. Studies had shown some relationship between AMD and mutations in the complement system of the innate immune response [13]. Genentech was not the only one that turned to the complement cascade as a therapeutic target [13]. Other molecules had been developed to inhibit complement factors C3 and C5 [13]. Lampalizumab, however, was the only molecule being developed that targeted Factor D, a complement factor upstream of C3 and C5 [13].

In what was characterized as a successful Phase II trial (MAHALO), a total of 129 patients were randomized to receive either lampalizumab 10 mg monthly or bimonthly, or sham treatments monthly or bimonthly [13]. The primary endpoint was mean change in GA area from baseline to 18 months [13]. The results showed a 20.4% reduction in mean change in GA area from

baseline in the lampalizumab monthly treatment group, which was statistically significant based on a prespecified significance level of $p < 0.2$ [13]. No benefit was observed in the bimonthly treatment group. In addition, a retrospective subgroup analysis was performed to assess the validity of certain exploratory biomarkers. Given the association of AMD with mutations in the complement system, the subgroup analysis focused on patients having mutations in certain complement system genes (CFH, C3, C2/CFB, and CFI) [2]. First, in the monthly treatment population, this small subgroup (n = 17) showed a "44% reduction [in GA enlargement] at the study's conclusion compared to the sham group" ($p < 0.005$) [13]. Of note, the percent reduction was stated in comparison to sham treatment, not baseline, as in the primary endpoint. Data for this subgroup at baseline and 18 months were not published. Second, in the sham population, patients with CFI mutations (n = 14) were found to have "significantly more atrophy" than patients without mutations [13]. Taken together, the investigators concluded that "the treatment response was magnified in patients with the complement factor I [CFI] biomarker," and that in the absence of treatment (sham group), CFI mutations caused accelerated disease progression [13].

Based on the results of the MAHALO trial and subgroup analysis, Genentech proceeded into phase III with two identically designed clinical trials, CHROMA and SPECTRI. This time, lampalizumab 10 mg was administered either every 4 weeks or every 6 weeks (likely given the lack of efficacy with bimonthly dosing seen in MAHALO) [14]. These were compared with sham treatments administered every 4 weeks or every 6 weeks [14]. The studies also prospectively investigated CFI profile as a genetic biomarker, stratifying and randomizing patients according to their CFI profile status [14]. The studies were designed to run up to 92

weeks but, following primary analysis of SPECTRI data at week 48, both were terminated early at the sponsor's recommendation due to lack of efficacy [14]. No benefit of lampalizumab over sham was observed, including in the CFI+ subgroup as well as other prespecified subgroups [14].

What can be said about the failure of lampalizumab so late in its clinical development?

Less Stringent Trial Design

First, with a prespecified $p < 0.2$, there would have been as high as a 1 in 5 probability of the study falsely detecting a significant difference between lampalizumab and sham, compared to only a 1 in 20 probability had the more stringent and conventional value of $p < 0.05$ been used. A larger *p*-value does not automatically negate a positive finding, but does mean that there is a higher risk that the positive finding is false. Investigators may design studies with such a higher level of risk for a multitude of reasons. For example, in an early phase trial, a small number of subjects may mean that the study is not adequately powered to detect the anticipated treatment effect. It may also be that the anticipated treatment effect could be small or difficult to detect.

The MAHALO trial enrolled a relatively small sample size of 129 subjects, and suffered significant attrition (30 subjects), such that a modified intention-to-treat strategy (details not disclosed) was ultimately used to analyze the data for a total of 123 subjects. Further, there may have been uncertainty about the treatment effect, especially since lampalizumab was the only investigational compound targeting Factor D at the time. It may be for these reasons that the less

rigorous value of $p < 0.2$ was chosen. Whatever the case, it allowed for a higher likelihood of an erroneous result that would serve as the basis for a "go" to phase III decision. Indeed, as discussed above, lampalizumab did not reduce GA enlargement compared to sham after 48 weeks of treatment in the phase III CHROMA and SPECTRI trials.

Incomplete Understanding of Disease Pathway

After the failures of CHROMA and SPECTRI, the trial investigators posed the question "whether the complement cascade is an appropriate intraocular therapeutic target for GA, at least through the alternative pathway via complement factor D or downstream in the cascade via C5" [14]. They also noted numerous failures of other study drugs attempting to exploit the complement cascade. Essentially what they were saying was that the scientific community still lacked a thorough understanding of the GA disease pathway.

Ideally, disease etiology should be well understood by the time of phase III trials. Phase III failure of lampalizumab could possibly have been avoided had that been the case. Unfortunately, even as late in time as the phase II MAHALO trial, the complement cascade posed but "an intriguing clinical therapeutic target" [13]. It was not known (and is still not known) whether mutations in the complement cascade were causative or merely associative with GA, and thus whether targeting the complement cascade on some level could actually alter the course of GA.

Retrospective Subgroup Analysis and False Biomarker Identification

The subgroup analysis from the MAHALO trial appears to have identified a spurious biomarker, CFI profile status. While subgroup analyses can be useful for exploratory and hypothesis-generating purposes, their results should be viewed cautiously for a variety of well-recognized reasons: insufficient power, lack of stratified randomization, multiplicity, "data dredging," etc. [15, 16]. Large clinical trials, for example, are typically powered to detect a difference in the primary endpoint between treatment and control groups 80-90% of the time [16]. However, when it comes to detecting differences between subgroups, it has been estimated that power may be reduced to as low as 20-30%, leading more often than not to false positive and false negative findings [16]. Not only is the sample size comprising a subgroup smaller than the overall treatment population, detecting differences in treatment effect between two treatment subgroups is usually more difficult than detecting differences between treatment and placebo groups [16].

In MAHALO, trial investigators acknowledged that "the study was not powered to look at these differences [in CFI profile status], and they remain to be tested in larger populations." Indeed, the subgroup of CFI+ patients in the monthly treatment group numbered only 17. In contrast, in CHROMA and SPECTRI, patients were prospectively stratified according to CFI biomarker status, with a 95% power to detect a difference in mean GA area in CFI+ patients, and 80% power in CFI- patients (due to the 3:2 ratio of CFI+ to CFI- patients). The fact that such drastically different results were seen in MAHALO (44% reduction in GA area enlargement in CFI+ patients compared to sham) versus CHROMA and SPECTRI (no difference between CFI+ and CFI- patients) underscores how lack of power in subgroup analyses can easily lead to false positives, as here, as well as false negatives. The investigators themselves commented that "the much larger prospective analysis of CHROMA and of SPECTRI" did not validate CFI profile

status as a biomarker for GA progression [14]. And in fact, other subsequent studies also did not replicate the MAHALO findings [14]. The trial investigators ultimately retracted the MAHALO publication, stating that the results of the trial represented only "preliminary analyses" [17].

Onartuzumab and Rilotumumab. The late stage failures of onartuzumab (developed by Roche) and rilotumumab (developed by Amgen) are also attributable, at least in part, to posthoc subgroup analyses and false biomarker identification from phase II that were prospectively tested for the first time in large phase III trials. In both cases, MET-positive tumor status was thought to be a biomarker predictive of enhanced treatment response, based on phase II subgroup analyses [18, 19]. While the "exploratory nature of the analysis" and the fact that "MET expression analysis was not prospectively designed" were recognized as major limitations of the phase II trials, in both cases the decisions were made to conduct phase III trials in patients with MET-positive tumors [18, 19]. Perhaps unsurprisingly, the phase III trials for onartuzumab in non-small cell lung cancer (METLUNG) and gastroesophogeal cancer (METGASTRIC) and the phase III trial of rilotumumab in gastroesophageal cancer (RILOMET-1) failed to confirm phase II findings [20, 21, 22]. In fact, METGASTRIC was stopped early due to lack of efficacy and RILOMET-1 was stopped early due to increased deaths in the rilotumumab arm [21, 22]. Therefore, not only can retrospective subgroup analysis falsely identify a biomarker of enhanced efficacy, it can fail to flag increased risk of drug safety events in the biomarker-positive subgroup.

We next turn to bococizumab, which suffered late stage failure during its clinical development due to formation of antidrug antibodies, sub-optimal dosing, and multiple simultaneous clinical

trials. In addition, tralokinumab is discussed as a variation on the theme of consequences from running multiple simultaneous clinical trials.

Bococizumab

Reduction of plasma low-density liproprotein cholesterol (LDL, LDL-C) is associated with decreased risk of cardiovascular events such as stroke and myocardial infarction. Statin therapy is a mainstay of cardiovascular risk reduction, however, individual response to statins can vary considerably and include intolerance and resistance.

Recently, inhibitors of proprotein convertase subtilisin-kexin type 9 (PCSK9) have been developed as an alternative to statins for reducing LDL and, as a result, cardiovascular risk. PCSK9 binding to the LDL receptor induces degradation of the LDL receptor. Inhibition of PCSK9 therefore results in greater abundance of LDL receptors on hepatocyte cell surfaces, allowing for increased removal of LDLs from circulation and decreased plasma LDL levels.

Bococizumab is a humanized monoclonal antibody developed by Pfizer that targets PCSK9 [23]. It consists of approximately 3% murine sequence, located at the complementarity-determining regions (CDR) [23]. At the time of development, at least two other PCSK9 inhibitors were also in development, alirocumab (Sanofi and Regeneron) and evolocumab (Amgen) [23]. Both of these were fully human monoclonal antibodies [23].

In a phase II dose-ranging study, the LDL lowering effect of bococizumab was tested at three different biweekly doses and two different monthly doses [24]. The 150 mg biweekly dose produced the greatest LDL lowering effect and was selected as the dose for phase III testing [23, 25]. Regarding safety, antidrug antibodies (ADAs) were detected in 18 of 251 (7%) patients who received bococizumab, however, "[t]he AEs reported in subjects with ADAs were similar to those observed in subjects without ADAs, with no signs or symptoms of hypersensitivity associated with positive ADA titers" [24]. Additionally, with the exception of a single patient whose LDL levels trended toward baseline, "little variation was found in LDL-C response between subjects with and without ADAs" [24].

On the basis of the phase II results, bococizumab was tested in a series of six parallel, multicenter, randomized, placebo-controlled phase III trials (collectively, SPIRE) to confirm efficacy in LDL reduction [25]. Simultaneously, two other parallel, multicenter, randomized, placebo-controlled phase III trials (SPIRE-1 and SPIRE-2) were underway to evaluate whether bococizumab, in lowering LDL levels, actually reduced the incidence of major adverse cardiovascular events [23].

In the SPIRE trials, bococizumab confirmed efficacy in significantly lowering LDL levels compared to placebo, with a mean percent change of -54.2% from baseline at 12 weeks [25]. However, 48% of patients who received bococizumab developed ADAs by year 1, with ADAs detectable starting in week 12 in most cases [25]. Moreover, 29% of patients who received bococizumab developed neutralizing antibodies [25]. ADA formation, and in particular the neutralizing antibodies, attenuated the LDL lowering effect of bococizumab over time, with

greater attenuation seen in patients with higher ADA titers [25]. Further, in both patients with and without positive ADA titers, wide variation was observed in LDL lowering effect [25].

On November 1, 2016, Pfizer terminated the bococizumab development program, citing immunogenicity and variability in response to the study drug [26]. By this point in time, alirocumab (Praluent®) and evolocumab (Repatha®) had already been approved by the FDA, although initial sales were sluggish [27, 28].

Following Pfizer's announcement, the SPIRE-1 and SPIRE-2 cardiovascular outcomes trials were stopped early [23]. At the time of stopping, SPIRE-1, which evaluated bococizumab in lower risk patients over a shorter period of time (median 7 months), showed no benefit of bococizumab over placebo in reducing major adverse cardiovascular events [23]. SPIRE-2 evaluated higher risk patients over a longer period of time (median 12 months), and actually showed significant reduction in adverse cardiovascular outcomes compared to placebo (hazard ratio 0.79; 95% CI, 0.65 to 0.97; P = 0.02) [23]. ADA formation and variability in response to bococizumab were not discussed for SPIRE-1 and SPIRE-2 in the publication that followed [23].

What led to the demise of bococizumab in phase III clinical trials?

Antidrug Antibodies

ADAs, in this case, spelled the doom of bococizumab, both as a safety and an efficacy-attenuating issue. Of course, a humanized monoclonal antibody like bococizumab carries an

intrinsically higher risk of immunogenicity than a fully human monoclonal antibody like alirocumab (Praluent®) and evolocumab (Repatha®). However, it would be an oversimplification to suggest that the failure of bococizumab teaches us to avoid anything but fully human monoclonal antibodies in drug development. The plethora of approved safe and effective humanized and chimeric monoclonal antibodies on the market—rituximab, infliximab, pembrolizumab, ocrelizumab, to name a few—immediately discredit such a notion. Moreover, even fully human monoclonal antibodies such as adalimumab and golimumab can themselves induce immunogenic reactions ranging widely between "negligible, tolerable, and marked" [29].

The more relevant question is why the immunogenicity issue with bococizumab emerged so late in clinical development.

First, as safety issues, adverse events are not primary endpoints that studies are usually powered to detect, or at least detect at an accurate magnitude. In the bococizumab phase II study, ADA formation was in fact detected, but given the small treatment population (only 251 patients received bococizumab), the 7% of patients who developed ADAs could have been much higher or, as it turned out, much lower than the true incidence. Thus, it appears that the phase II study yielded a false low as to an important safety signal. And while trial investigators will typically scrutinize safety data regardless of statistical significance, unlike efficacy, here it appears that the investigators were perhaps overly optimistic about the 7% incidence of ADA formation with bococizumab, noting that "[r]ecent clinical trials of other PCSK9 inhibitors [alirocumab] have found ADA incidences of up to 11.5% in both phase 2 and phase 3 studies" [24]. However, they did note that the lack of standardized ADA testing methods made direct comparison difficult,

and that a “more comprehensive assessment” of ADA formation would be undertaken in Phase III testing [24].

Sub-Optimal Dosing

Bococizumab dose reductions occurred in the phase II trial when LDL levels fell below a prespecified cutoff (≤25 mg/dL) [24]. The reason for allowing dose reductions was that the safety of very low LDL levels, based on the existing literature, was unclear and possibly detrimental [24]. For the 46 patients in the 150 mg biweekly dosing regimen (which was the dosing regimen eventually selected for phase III trials), only 61% (28) of them were receiving the same dose at the end of the trial [19]. The other 39% (18) of patients were reduced to lower biweekly doses of 100 mg, 50 mg, or even 25 mg [24].

While these data speak to the dramatic LDL lowering effects of bococizumab, the biweekly 150 mg dose, while producing the greatest LDL lowering effect, may not have been the optimal dose to select for phase III trials, as it necessitated multiple, serial dose reductions in a large proportion of patients. As the investigators noted, reduced LDL levels were maintained “despite protocol-stipulated dose reductions in a large proportion of subjects” [24], suggesting that reduced doses were still efficacious and that the 150 mg dose might have produced more efficacy than was needed or desired, given the unknown danger of extremely low LDL levels. Further, it is possible that a consequence of the dose reductions is that it muted the magnitude of the ADA issue during phase II. Had the ADA issue fully manifested in phase II, the time and cost of running 8 phase III trials might have been avoided.

Multiple Simultaneous Clinical Trials

A final point that can be made is a caveat against running multiple clinical trials simultaneously, instead of applying lessons learned from one clinical trial to the next [30]. With bococizumab, Pfizer ambitiously chose to run 8 simultaneous multicenter phase III clinical trials: the 6 smaller trials comprising SPIRE and the larger SPIRE-1 and SPIRE-2 trials that had a combined enrollment of 27,438 patients. Certainly, there was pressure to bring a new PCSK9 drug to market, especially with alirocumab (Praluent®) and evolocumab (Repatha®) recently approved. However, market forces are ever present. Had Pfizer waited for the SPIRE results before initiating SPIRE-1 and SPIRE-2, the latter two trials would never have gotten underway, and the tremendous time and cost of these trials could have been avoided entirely.

Tralokinumab. A similar error was committed in the case of tralokinumab, an IL-13 antagonist developed by AstraZeneca for the treatment of asthma. In phase III clinical testing of tralokinumab, rather than prespecifying biomarker subgroups, the plan was to identify biomarkers in the first trial (STRATOS-1) that would then be investigated in the second trial (STRATOS-2) [31]. STRATOS-1 identified fractional exhaled nitric oxide (FENO) as the putative biomarker best predictive of enhanced benefit to tralokinumab [31]. Once identified, the protocol for STRATOS-2 was amended so that FENO-high patients became the study population for the primary endpoint, while the original all-comers population became the secondary study population [31].

The problem was that despite staggering the design of the study protocols, the trials themselves were run simultaneously [31]. As such, the prevalence of FENO-high participants in STRATOS-2 was basically "consistent with STRATOS-1" [31]. As the trial investigators acknowledged, because "the trials proceeded at the same time, there was no opportunity to enrich the STRATOS-2 population for a FENO-high subgroup" [31]. This "could therefore have affected the power of STRATOS 2" [31]. It is unclear why the two trials were carried out in parallel, as identification of a promising biomarker in STRATOS-1 obviously would not automatically enrich STRATOS-2 with biomarker-positive patients. In this case, by carrying out the two trials simultaneously instead of sequentially, the trial investigators defeated the purpose of their original trial design. Identification of a putative biomarker was made in vain because investigators were not able to effectively test the biomarker in a biomarker-positive enriched population. In the end, STRATOS-1 and STRATOS-2 were essentially no more than two identically designed clinical trials carried out in a general, all-comers population.

We turn next to tabalumab, a promising candidate with an expansive development program across a variety of indications. It ultimately failed due to sub-optimal dosing and route of administration, as well as bypassing phase II trials and proceeding directly from phase I to phase III. Unfortunately, as we shall see, these errors were entirely avoidable but had devastating consequences for the tabalumab development program.

Tabalumab

B-cell dysfunction has been observed in a number of autoimmune diseases, including systemic lupus erythematosus (SLE) and rheumatoid arthritis (RA). Inhibition of B-cell function has therefore become of great interest in drug discovery efforts, including, in particular, inhibition of B-cell activating factor (BAFF). BAFF is a member of the tumor necrosis factor (TNF) superfamily and is involved in B-cell development, survival, and differentiation. It exists in membrane-bound and soluble forms.

Tabalumab is a fully human monoclonal antibody developed by Eli Lilly that binds to and neutralizes both membrane-bound and soluble BAFF [32]. During the course of its development, the FDA approved belimumab (Benlysta®), another anti-BAFF monoclonal antibody developed by GSK, for the treatment of SLE [32]. Belimumab binds only to soluble BAFF [32].

Eli Lilly sought to develop tabalumab for a multiplicity of indications, including RA, SLE, multiple myeloma (MM), multiple sclerosis (MS), and end stage renal disease (ESRD) patients awaiting transplant and requiring immunosuppression. As summarized in Table 2 below, a total of 23 tabalumab clinical trials were registered at www.clinicaltrials.gov or listed in PubMed under a search for "tabalumab" (clinical trial registries of the EU and WHO were also searched but did not yield additional results).

Our discussion will focus on tabalumab for SLE, with reference to the tabalumab development program for RA, which was progressing in tandem. Results first became available for the phase III clinical trials for RA (FLEX-O, FLEX-V, and FLEX-M). Tabalumab failed to demonstrate

efficacy in reducing RA disease activity in all three trials, and the FLEX-V and FLEX-M trials were in fact terminated early [34, 35]. Shortly thereafter, Eli Lilly announced that it would discontinue the tabalumab development program for RA, while continuing with the SLE and multiple myeloma development programs [36].

[Table 2 goes here.]

Results of the phase III clinical trials for SLE (ILLUMINATE-1, ILLUMINATE-2) became available next. In ILLUMINATE-1, tabalumab again failed to meet primary efficacy endpoints in both the biweekly and monthly dosing regimens that were studied [32]. In ILLUMINATE-2, efficacy of biweekly dosing of tabalumab reached statistical significance, but not monthly dosing [33]. In light of these results, Eli Lilly announced that it was terminating the entire tabalumab development program [37].

What went wrong with tabalumab for SLE? Why did it fail in phase III clinical trials for SLE, after already failing in phase III trials for RA?

Sub-Optimal Dose and Route of Administration

In the phase III ILLUMINATE trials, the dose and route of administration for tabalumab was tested in SLE patients for apparently the first time. Earlier, in a combined RA/SLE phase I study, a single dose of tabalumab was administered intravenously to a total of 5 patients with SLE (and 17 patients with RA) [38]. In the phase III ILLUMINATE trials, subcutaneous dosing

was selected instead [32, 33]. It is possible that the decision to proceed with subcutaneous dosing was based on the multiple phase II trials that were conducted for RA. Intravenous dosing had shown mixed results in two phase II trials for RA, with one (NCT00308282) demonstrating efficacy [39] and another (NCT00689728) lack of efficacy [40]. On the other hand, subcutaneous administration proved efficacious in two other phase II trials for RA (NCT00785928, NCT00837811) [41 42], and ultimately the subcutaneous route was selected for the phase III clinical trials for RA [34, 35]. Therefore, it could have been considered expeditious to do the same for phase III in SLE as well.

The selected dose for phase III, however, is less easily surmised. As discussed, in phase I a total of 5 SLE patients received single doses of intravenous tabalumab, either 0.125 mg/kg or 2.0 mg/kg [38]. In the phase III ILLUMINATE trials, subjects were given an initial subcutaneous loading dose of 240 mg, followed by 120 mg subcutaneously either biweekly or monthly [32, 33]. There must have been a basis for the doses selected, but in the absence of any phase II or other study in SLE, the only other source of data would have been from the clinical trials for other indications. At least 3 phase II trials were run for RA utilizing subcutaneous dosing, as well as other phase II trials for MM, MS, and ESRD, but none for SLE. The 120 mg monthly dosing from the phase III ILLLUMINATE trials was previously used in two of the RA phase II trials [41, 42], while the 240 mg loading dose found precedent in an ESRD phase II trial [43]. No apparent precedent existed, however, for the 120 mg biweekly dosing used in the ILLUMINATE trials, and yet as it turned out, it was the only dosing regimen that demonstrated any significant efficacy [33].

It was unfortunate that an efficacious dosing regimen for tabalumab was only discovered in phase III. That dose selection for phase III in SLE should have been based on doses previously studied in patients with SLE, rather than other diseases, is an issue the investigators themselves implicitly acknowledge. In discussing the ILLUMINATE study results, in light of the fact that the more frequent biweekly dosing met the primary efficacy endpoint in ILLUMINATE-2, trial investigators stated that "it remains possible that higher doses or more frequent dosing may lead to better efficacy" [32], but unfortunately "the optimal dose remains unknown either for the cross-section of international patients who were studied or for individual patients" [33]. These are the types of questions that could have been explored in a phase II dose-ranging study in SLE patients. Reasonable certainty of an efficacious dosing regimen based on a phase II dose-ranging study could have led to a successful phase III trial and subsequent BLA submission. Instead, tabalumab failed in phase III trials and the entire tabalumab development program was shut down. The importance of optimal dosing in clinical trials cannot be understated.

Skipping Phase II Testing and Proceeding Directly from Phase I to Phase III

It is certainly remarkable that there was no efficacy data for tabalumab in SLE when Eli Lilly proceeded into phase III trials, only early safety and PK data from single doses in phase I. Eli Lilly commenced two large-scale phase III trials without any actual idea of the study drug's efficacy in SLE. Yet it conducted phase II trials for every other indication it pursued: RA, MM, MS, and ESRD.

One can speculate regarding the decision to proceed immediately to phase III for SLE. It is possible that the plethora of phase II trials in other disease states was viewed as having provided sufficiently generalizable information for tabalumab. As discussed above, it certainly appears that dosing for the phase III ILLUMINATE trials was borrowed from the other phase II trials. In addition, as was the case with bococizumab, there was perhaps a sense of intense market pressure. Before the ILLUMINATE trials got underway, GSK had already received FDA approval for belimumab (Benlysta®) as an intravenous formulation and was in the process of seeking approval for a self-injectable subcutaneous formulation [44], like tabalumab. Whatever the reasons, in the end the demise of tabalumab reminds us that cutting corners in drug development often leads to avoidable late stage failures.

We continue our discussion with solanezumab, a failed drug candidate in the notoriously challenging field of Alzheimer's disease research. The late stage failure of solanezumab can be attributed to improper inclusion criteria in phase III, clinical data that did not accord with preclinical results, and failure of target engagement to modify disease. With respect to the latter, lebrikizumab and tralokinumab are also discussed as additional examples.

Solanezumab

Drug failures in Alzheimer's disease (AD) are all too common. With a 99.6% failure rate, a total of 5 FDA approved drugs for AD (symptom-modifying only), and no FDA approvals for new AD agents since 2003, the path to finding a cure for AD has been difficult indeed [45]. Solanezumab was another promising drug candidate for AD that ultimately failed—multiple times—late in its clinical development.

Developed by Eli Lilly, solanezumab is a humanized monoclonal antibody that binds to and increases clearance of soluble amyloid beta protein. Beta amyloid plaques are considered a hallmark of AD, and they have been a target of AD drug development efforts for over 20 years.

Preclinical testing of m266.2, the murine analog of solanezumab, demonstrated dose-dependent increases in plasma beta amyloid in transgenic PDAPP mice, a mouse model of AD [46, 47]. Investigators therefore posited a "peripheral sink hypothesis," whereby binding of m266.2 to soluble plasma beta amyloid caused efflux of beta amyloid from CSF to plasma, increasing plasma concentrations and ultimately clearance [46]. m266.2 was also detected in CSF at approximately 0.1% of plasma levels [46]. Though CSF penetration was low, investigators observed increases in CSF beta amyloid as well, and hypothesized that m266.2 acted in the same manner in CSF as plasma, breaking down beta amyloid plaques in the brain [46]. Indeed, beta amyloid deposition was significantly reduced in mice treated with m266.2, compared with IgG control [46].

Subsequently, in the phase I, phase II, and phase III (EXPEDITION1, EXPEDITION2, and EXPEDITION3) trials, increases in total plasma and CSF concentrations were similarly observed in human subjects receiving solanezumab [48, 49, 50, 51]. However, a reduction in beta amyloid plaque burden, as assessed by PET imaging, was not seen in these patients in the phase III trials (imaging was not performed in the phase I and II studies) [50, 51]. Moreover, in the identically designed EXPEDITION1 and EXPEDITION2 phase III trials, solanezumab did not reduce the decline in cognition or function in patients with mild to moderate AD compared to placebo [50].

In a prespecified subgroup analysis of mild AD patients, based on pooled data from both EXPEDITION1 and EXPEDITION2, less cognitive decline was seen in the solanezumab group compared to placebo [52]. Based on this pooled secondary analysis, a third phase III trial, EXPEDITION3, was undertaken to assess the efficacy of solanezumab in patients with mild AD [51]. Unfortunately, EXPEDITION3 also failed to meet its primary efficacy endpoint, reduction in cognitive decline as measured by the ADAS-cog14 scale at week 80 [51].

Again we ask: what caused the failure of solanezumab so late in its clinical development?

Trial Design: Improper Inclusion Criteria

As previously described, solanezumb targets and binds beta amyloid protein. However, in EXPEDITION1 and EXPEDITION2, patients were not screened for the presence of beta amyloid plaques prior to study entry [50]. Instead, mild to moderate AD patients were included in the studies based on MMSE score and NINCDS-ADRDA criteria, neither of which takes beta amyloid accumulation into consideration [50]. For the mild AD subgroup, the trial investigators estimated that 25% of these patients did not have amyloid-related disease (*i.e.*, they likely had dementia without AD) and thus "would not have been expected to have a response to a treatment targeting Aβ [beta amyloid]" [51]. In EXPEDITION3, only patients with actual evidence of amyloid-related disease (by PET scan or CSF measurements) were enrolled in the study, although the study still failed to meet its primary endpoint [51]. Nevertheless, the lesson can be learned: properly defining the inclusion criteria for a study population can make the difference

between failure and success of a clinical trial. Whatever the study drug, the treatment population must actually exhibit the characteristic or trait that is targeted by the drug in order for there to be any chance of producing a treatment effect.

Preclinical Results Do Not Always Translate to Clinical Outcomes in Humans

Solanezumab is a reminder that preclinical results are not always reproduced in human studies. There are many reasons why animal models may fail to translate into outcomes in humans. Obviously, differences in biology and disease pathology and complexity exist between humans and any animal model. More pernicious, issues in trial design and reporting biases have been shown to plague many animal studies. For example, one review of 290 animal studies showed that lack of randomization and blinding resulted in a 3.4- and 3.2-fold increase, respectively, in claiming statistically significant results, compared to randomized and blinded studies [53]. Similarly, in terms of reporting bias, another study showed that preclinical studies with smaller numbers of animals tended to report greater treatment effect than larger animal studies [53]. These issues seriously undermine the value of animal studies in predicting clinical benefit in humans.

Interestingly, investigators in the solanezumab trials did not offer explanations as to why results of the animal studies (decreased beta amyloid plaque burden) were not replicated in the human studies. Instead, they appear to have abandoned their peripheral sink hypothesis to fit the latest results:

> "The administered dose of solanezumab did not reduce the burden of fibrillar amyloid, as assessed by means of florbetapir PET imaging. A lack of an effect on preexisting amyloid plaques in this trial is consistent with the results from earlier clinical data and nonclinical studies in animals. Antibodies that bind soluble Aβ [like solanezumab] would be expected to have only marginal effects on preexisting amyloid" [51].

It is as if the investigators never expected to see a decrease in beta amyloid plaque burden all along. And yet if that were the case, there would not have been much point in prescreening for AD patients with evidence of beta amyloid plaque burden and performing PET imaging to assess for changes in plaque burden.

Without the benefit of access to unpublished animal data, it is impossible to know what issues (lack of randomization, reporting bias, etc.) could have impacted the disparity between the animal and human study results seen with solanezumab. However, if we are to learn from such mistakes, it is necessary to objectively and squarely look them in the face rather than switching hypotheses to fit the data, because the data may be the problem. Certainly, hypotheses and theories can change over time, but they should neither be adopted nor abandoned in a premature or unmethodical manner.

Target Engagement May Not Result in Disease Modification

Finally, and here the study investigators do openly discuss the possibility, it may be that beta amyloid is not causative of but merely associated with AD. After more than 20 years of targeting beta amyloid by various mechanisms without slowing disease progression, the beta amyloid theory of AD has been called into question in recent years [54, 55, 56]. As the solanezumab investigators acknowledge, "if amyloid is not the cause of the disease, solanezumab would not be expected to slow disease progression" [51]. They go on to suggest that "the amyloid hypothesis will need to be considered in the context of accruing results from this trial and other clinical trials of antiamyloid therapies" [51]. Ideally, a clear and confirmed understanding of disease pathology should exist by the time a drug candidate enters clinical trials. Otherwise, even high levels of target engagement, as was the case with solanezumab [51], will not alter the course of the disease, while administering the drug in clinical trials puts patients at risk of serious adverse events.

Lebrikizumab, Tralokinumab, and Dupilumab. Another example in which target engagement failed to result in disease modification is lebrikizumab, an anti-IL-13 monoclonal antibody developed by Genentech for the treatment of asthma. After successful phase II testing, the phase III trials (LAVOLTA-1 and LAVOLTA-2) produced inconsistent results in reducing the rate of asthma exacerbations [57, 58]. In analyzing the results, investigators commented that it was "possible that maximal blockade of interleukin-13 was achieved but remained insufficient to substantially affect clinical endpoints" [58].

As with lebrikizumab, phase III testing of tralokinumab, another IL-13 antagonist (developed by AstraZeneca), produced inconsistent results (STRATOS-1 and STRATOS-2) [31]. According to

the trial investigators, the findings “add[ed] to available evidence that interleukin-13 blockade alone is insufficient to reduce asthma exacerbations in people with severe, uncontrolled asthma” [31].

In contrast, dupilumab (Dupixent®) is a monoclonal antibody developed by Regeneron and Sanofi that targets not only IL-13 but also IL-4. In phase II and III clinical trials, it was shown to be effective in reducing asthma exacerbations, sparing corticosteroid use, and improving lung function and quality of life [59, 60]. Dupilumab was approved by the FDA for moderate to severe asthma in October of 2018 [61]. This is not to suggest that hitting two targets instead of one is always the answer, but does underscore the fact that diseases often involve multiple complex signaling pathways, and blocking a single pathway may not be sufficient to produce disease modification.

We next examine the late stage failure of bavituximab, which can be attributed to poor quality phase II data as well difficult strategic and commercial factors that are not uncommon for smaller, early stage companies.

Bavituximab

Immunotherapy has proven to be an efficacious cancer treatment strategy in recent years. Immune checkpoints such as programmed cell death protein 1 (PD1), cytotoxic T-lymphocyte antigen 4 (CTLA4), and phospholipid phosphatidylserine (PS) are parts of normal immune system function, negatively regulating immune and inflammatory responses. Drugs that act as

checkpoint inhibitors, such as pembrolizumab (anti-PD1) and ipilimumab (anti-CTLA4), have been shown to lift negative regulation, activating the immune system against cancer cells.

Developed by Peregrine Pharmaceuticals, bavituximab is a chimeric monoclonal antibody that acts as a checkpoint inhibitor against PS on solid tumor cell surfaces. In preclinical models, bavituximab was shown to inhibit tumor growth, prolong survival, and synergistically enhance the efficacy of concomitant chemotherapy and radiation [62]. Following early phase safety and pharmacokinetic testing, a number of phase II trials were carried out in prostate cancer (NCT01335204), pancreatic cancer (NCT01272791), breast cancer (NCT00669591), advanced hepatocellular carcinoma (NCT01264705), non-small-cell lung cancer (NSCLC) (NCT01160601, NCT01138163), and hepatitis C (NCT01273948).

The only phase III trial of bavituximab to date is the SUNRISE trial, which compared the efficacy of docetaxel plus bavituximab or placebo in patients with NSCLC [62]. The primary endpoint was overall survival (OS) [62]. At the first planned interim analysis, however, an independent data monitoring committee recommended early stopping due to futility [62]. At the time, median OS in the docetaxel/bavituximab arm was 10.5 months, worse than the 10.9 month median OS seen in the docetaxel/placebo arm (HR 1.06; 95% CI 0.88-1.29; $p = 0.533$) [62]. There was also no significant difference seen in progression-free survival (PFS), one of the trial's secondary endpoints (HR 1.00; 95% CI 0.82-1.22; $p = 0.99$) [62].

What can account for such disappointing results?

Poor Quality Data from a Botched Phase II Trial

The late stage failure of bavituximab can be traced in part to problems originating in the preceding phase II NSCLC study. That study was designed as a 3-arm study, with patients randomized to receive docetaxel plus placebo, docetaxel plus lower dose bavituximab (1 mg/kg), or docetaxel plus higher dose bavituximab (3 mg/kg) [63]. The primary endpoint was overall response rate (ORR), and secondary endpoints included OS and PFS [63]. Due to a labeling discrepancy, some patients in the placebo arm received doses of bavituximab 1 mg/kg, and some patients in the bavituximab 1 mg/kg arm received doses of placebo [63]. Patients in the bavituximab 3 mg/kg arm were not affected by the labeling error [63]. Data from the placebo and bavituximab 1 mg/kg arms were pooled into a "combined control group," and the efficacy analyses were reported as "exploratory" [63].

As the labeling discrepancy did not impact the bavituximab 3 mg/kg arm, ORR in this treatment group could be considered a true standalone data point. Unfortunately, the observed 17.1% ORR (95% CI, 5.6%-28.6%, $p = 0.37$) fell short of the prespecified ORR of 26% required to achieve the primary endpoint [63]. As for the secondary endpoints, comparison of bavituximab 3 mg/kg against the combined control group was obviously problematic due to lack of a true control, nevertheless, significant differences in PFS and OS were not achieved. Median PFS was 4.5 months and 3.3 months in the bavituximab 3 mg/kg and combined control group, respectively (HR 0.74, 95% CI, 0.45-1.21, $p = 0.24$) [63]. Similarly, OS was 11.7 months and 7.3 months in the bavituximab 3 mg/kg and combined control group, respectively (HR 0.55, 95% CI, 0.40-1.10, $p = 0.11$) [63].

Certainly, advancing into phase III based on a phase II trial in which the primary endpoint was not met and "exploratory" secondary endpoints showed only non-significant trends was asking for trouble. The trial investigators themselves admitted that it was "not possible to draw firm conclusions" from the phase II trial [63], and yet the "go" to phase III decision was made.

One possible reason is that the phase II NSCLC study results were the best data they had. As discussed above, in addition to NSCLC, Peregrine tested bavituximab in numerous phase II trials across a variety of potential indications, including prostate cancer, pancreatic cancer, breast cancer, advanced hepatocellular carcinoma, and hepatitis C. As it happens, none of those study results were published, but while publication bias is a reality, negative results should not automatically be inferred from unpublished studies. In this case, however, the unpublished phase II studies, combined with the fact that phase III testing was not pursued in any disease state other than NSCLC, is somewhat more telling. The SUNRISE trial in NSCLC is the only phase III trial of bavituximab to date, and it may have been initiated because the botched phase II study produced the most positive data for bavituximab across all potential indications.

Strategic and Commercial Factors Pushed Bavituximab into Phase III

Strategic and commercial factors often drive decision-making in clinical development, and they appear to have exerted a strong force in the case of bavituximab. Bavituximab was one of only three molecules in Peregrine's portfolio, and the one furthest along in clinical development. Cotara®, a radiolabeled monoclonal antibody developed for brain cancer therapy, completed phase II testing but stalled in search of a partner to fund phase III development [64]. ^{124}I-

PGN650, a PS-targeting monoclonal antibody (like bavituximab) joined to radioisotope iodine-124 for use in PET imaging, also did not appear promising. A small phase 0 study had found that "tumor uptake [of ^{124}I-PGN650] was quite low and not sufficient for clinical studies" [65], and a phase I trial of ^{124}I-PGN650 was terminated early by the sponsor. No further clinical trials have been registered for ^{124}I-PGN650.

In this case, a slim pipeline and inadequate funding may have constrained Peregrine to pursue development of bavituximab for lack of any better options. If they terminated the bavituximab development program, they would have lost their only drug candidate with any real promise. It was, as viewed by some, a "make or break" situation for Peregrine [66]. Indeed, after disclosing that the data from the phase II NSCLC study were unreliable due to labeling errors, Peregrine's stock price plunged over 70% [67]. Clearly, the future of the company hinged almost entirely on the success or failure of bavituximab. The situation was perilous and Peregrine may have proceeded to phase III with bavituximab, despite foreseeable failure from the phase II results, because not doing so would have spelled the immediate end not only for bavituximab, but also Peregrine.

Subsequent events would appear to confirm how crucially Peregrine's fate was tied to the success or failure of bavituximab. In the wake of the phase III SUNSET trial, Peregrine sold the bavituximab program to Oncologie, Inc. [68], changed its name to Avid Bioservices [58], and transitioned from clinical stage biopharmaceutical company to a dedicated contract development and manufacturing organization (CDMO) [69].

The path of bavituximab and Peregrine Pharmaceuticals was a difficult one. In reviewing its phase II NSCLC study results and deciding whether to proceed to phase III, Peregrine faced either probable late stage failure of its lead molecule or probable failure of the company. Of course, not all companies, especially smaller ones, can boast a diversified and robust pipeline, as well as adequate funding for expensive clinical stage development. However, pursuing a drug candidate for lack of other better options is not an ideal recipe for success. The danger of a company positioning themselves in a situation like Peregrine with bavituximab is that instead of running a ruthlessly efficient clinical development program with a "fail early, fail cheap" mentality, the model becomes "fail late and delay the inevitable." In such a case, in addition to the actual cost of clinical development, there are the foregone opportunity costs that could have been allocated to developing more promising therapies. Further, ethical questions arise from continuing to expose patients to risk of harm with little to no expected benefit. Of course, taking these considerations into account requires an organization putting the greater good above its own. Admittedly, that is no easy task.

Lastly, we turn our attention to sirukumab, which is unique among the late stage failures we have surveyed in that it proceeded beyond phase III clinical trials, but still ultimately failed to obtain regulatory approval. As we shall see, the FDA viewed the totality of the clinical data differently than the sponsor. In addition, provisions regarding escape in the clinical trials likely created a bias against sirukumab in terms of the efficacy data. We also briefly discuss tremelimumab as another example wherein escape provisions compromised the data of ultimate interest.

Sirukumab

Rheumatoid arthritis (RA) is a chronic, progressive inflammatory disease that causes joint destruction, deformity, and physical debilitation. Elevated levels of interleukin-6 (IL-6) have been correlated with disease activity in RA patients, making it an attractive target for RA therapies. Tocilizumab (Actemra®) and sarilumab (Kevzara®) are approved IL-6 inhibitors that target the IL-6 receptor. Sirukumab is a monoclonal antibody developed by Janssen Pharmaceuticals that targets the IL-6 cytokine.

Sirukumab successfully completed three pivotal phase III trials that confirmed superior efficacy to placebo (SIRROUND-D, SIRROUND-T) [70, 71] and noninferiority to adalimumab (Humira®), an TNF inhibitor approved for treating RA (SIRROUND-H) [72]. As far as safety, trial investigators reported that sirukumab had an "expected safety profile" [70] that "did not raise any new concerns and was consistent with those reported for agents targeting the IL-6 receptor, such as tocilizumab and sarilumab" [50, 71, 72]. Janssen submitted a biologics license application (BLA) for sirukumab for adults with moderate to severe RA and an inadequate response or intolerance to one or more disease-modifying anti-rheumatic drugs (DMARDs) [73].

Citing safety concerns, however, the FDA's Arthritis Advisory Committee voted 12-1 against approving sirukumab [74]. Specifically, they noted safety imbalances in death, serious adverse events, MACE, serious infection, and malignancy, as well as increased risk of GI perforation and laboratory abnormalities with sirukumab [75]. Ultimately, the FDA issued a complete response letter to Janssen, indicating that additional data were needed to evaluate the safety profile of

sirukumab [76]. Janssen subsequently announced that it was terminating the sirukumab program and would "prioritize other assets in [its] portfolio" [77].

In a sense, sirukumab is a unique case study in that it represents an ultra-late stage failure, beyond phase III and in the BLA stage. How was it that Janssen and the FDA adopted such widely differing views on the safety of sirukumab? Did the clinical data actually reveal true safety signals, or were they the result of flawed trial design?

Examining the Totality of the Evidence from All Angles

The results of SIRROUND-D, SIRROUND-T, and SIRROUND-H were individually published, and in isolation, each of these pivotal phase III studies could be said to demonstrate an acceptable safety profile for sirukumab. In the SIRROUND-D study, for example, only 0.2% of patients in the placebo group experienced GI perforation, along with 0.2% in the sirukumab 50 mg q4w group, and 0% in the sirukumab 100 mg q2w group [70]. Comparably, in the SIRROUND-H study, incidence of GI perforation was 0%, 0.5%, and 0.5% in the adalimumab, sirukumab 50 mg q4w, and sirukumab 100 mg q2w groups, respectively [72]. In analyzing the results of each study on its own, trial investigators were led to conclude:

> "The safety profile of sirukumab did not raise any new concerns and was consistent with those reported for agents targeting the IL-6 receptor, such as tocilizumab and sarilumab. The proportions of patients experiencing AEs and SAEs were relatively similar

between treatment groups and the types of AEs and SAEs were similar through the 52-week study period" [70] (SIRROUND-D).

"The safety and tolerability profile in our study was similar to that for other drugs that target the interleukin-6 signalling pathway, including the anti-interleukin-6 receptor drugs tocilizumab and sarilumab. Gastrointestinal perforations were noted in this study and are a known risk of drugs that target interleukin-6 signalling. Laboratory abnormalities in this study with sirukumab are consistent with the safety profile of anti-interleukin-6 receptor drugs" [71] (SIRROUND-T).

"The safety profile of sirukumab was generally consistent with the known safety profile of anti-IL-6R antibody treatment and previous sirukumab RA studies" [72] (SIRROUND-H).

It is important, however, to examine the data from all angles and perspectives. In reviewing the BLA for sirukumab, the FDA considered data from SIRROUND-D, SIRROUND-T, and SIRROUND-H collectively, in addition to data from the long-term extension periods from these studies; a phase 2 dose ranging study; and an additional safety study in Japan [75]. In doing so they discovered disturbing safety imbalances in death, serious adverse events, MACE, serious infection, and malignancy. For example, a total of 35 deaths occurred in sirukumab-treated patients across all studies, compared to 1 death in the placebo-treated population [75]. 18

sirukumab-treated patients developed malignancies in the pooled SIRROUND-D and SIRROUND-T analysis, compared to 1 patient in the pooled placebo groups [75]. The FDA also noted increased risk of GI perforation and laboratory abnormalities (lipid elevations, neutropenia, thrombocytopenia, and elevated liver function enzymes) with sirukumab [75].

The results of the SIRROUND studies looked very different when interpreted individually by the drug sponsor and collectively by the FDA. Because a multitude of factors can impact the outcome of a clinical trial (non-adherence, concomitant medications, loss to follow-up, etc.), it is important not to accord undue weight to the result of any single trial, even if it is a randomized, double-blind, placebo-controlled trial. It is for this reason that the FDA historically interpreted "adequate and well-controlled investigations" in support of drug efficacy claims under § 505(d) of the Federal Food, Drug, and Cosmetic Act of 1962 (FFDCA) as requiring "at least two adequate and well-controlled studies, *each convincing on its own*" [78] (emphasis added). Better still to have data from multiple clinical trials available to compare and consolidate. While the 1997 Food and Drug Administration Modernization Act (FDAMA) amended § 505(d) of FFDCA to allow evidence of efficacy to rest on "one adequate and well-controlled clinical investigation and confirmatory evidence," such "confirmatory evidence" still commonly translates into a second phase III study with similar or identical trial design as the first "adequate and well-controlled study" [78]. In the end, whatever form the evidence takes, it must be examined at all levels to minimize bias in interpreting the data and clearly understand the true safety and efficacy profile of the drug.

Trial Design: Permitting Escape into the Sirukumab Treatment Arms

The SIRROUND-D and SIRROUND-T trials permitted patients in the placebo group with less than 20% improvement from baseline in swollen and tender joint count to escape to sirukumab treatment at prespecified timepoints [70, 71]. A large number of patients originally assigned to placebo in fact escaped to sirukumab—211 of 556 placebo patients in SIRROUND-D and 94 of 294 placebo patients in SIRROUND-T [70, 71]. Janssen suggested that because the trial design allowed for escape, the placebo groups were depleted of patients with high disease activity while the sirukumab groups were enriched with these patients, creating a "bias against sirukumab arms" for death and other adverse events [79]. The FDA acknowledged that it was "unclear whether the imbalance in all-cause mortality is a true safety signal or whether it is a result of bias due to the study design" [80]. One member of the Arthritis Advisory Committee commented that it was "a very close call" [74].

Although the possibility of escape mitigates the ethical challenge of administering of placebo to patients with active disease, it need not compromise data in the treatment arms of a study. In the SIRROUND trials, alternative trial designs could have avoided enriching the sirukumab arms with high disease activity patients. For example, patients escaping from placebo could have been treated with standard of care RA therapies instead of sirukumab. While the study populations generally consisted of patients refractory to one or more conventional therapies, some of whom continued on conventional therapies during the trials, many treatment options were available for RA [82, 83]. In addition to switching agents, dosing and frequency adjustments could have been made (as was done with escape patients in the adalimumab active comparator arm in

SIRROUND-H [72]), and combination therapy with these agents could have been attempted according to treatment guidelines [83]. Short-term, low-dose corticosteroid therapy could also have been administered according to treatment guidelines [84].

In such a close case as this, structuring the escape mechanism so as to preserve the integrity of the safety data in the sirukumab arms could have resulted in clean data in support of BLA approval. Instead, Janssen was left with inconclusive data that triggered a complete response letter requesting additional clinical studies to further evaluate the safety profile of sirukumab [76] (that Janssen was not willing to undertake [77]). The lesson here is to not lose sight of the ultimate goal when designing a clinical trial and to ensure that provisions regarding escape and other details do not compromise the data of ultimate interest.

Tremelimumab. A similar situation occurred in the case of tremelimumab, an anti-CTLA-4 monoclonal antibody developed by AstraZeneca for multiple indications, including non-small cell lung cancer, head and neck squamous cell carcinoma, and metastatic melanoma. It failed phase III testing in all of these indications [85, 86, 87]. In the phase III trial for metastatic melanoma (the only indication to date with published trial data), patients were randomly assigned to receive either tremelimumab or standard of care chemotherapy (temozolomide or dacarbazine) [87]. In this case, the trial protocol did not allow patients in the chemotherapy arm to escape or cross over to the tremelimumab arm [87]. However, in what was effectively an alternative escape mechanism, at least 14% of the patients in the chemotherapy arm (an underestimate, according to trial investigators) reported receiving ipilimumab (Yervoy®), an

FDA-approved monoclonal antibody developed by Bristol Myers Squibb that also targets CTLA-4, like tremelimumab [87].

In analyzing the late-stage failure of tremelimumab, investigators suggested that "[u]se of CTLA4 blockade in both arms of this study could have decreased the power of the study to demonstrate a statistically significant difference in survival and biased the estimates of survival in the control arm" [87]. As discussed above, preserving the data of ultimate interest from confounding or dilution is critical to designing and executing a successful clinical trial. Interestingly, in the phase III studies of ipilimumab, patients who had previously received treatment with any anti-CTLA4 antibody were excluded, and during the trial, patients in the control groups were not permitted to escape into the ipilimumab arms, nor were they allowed to receive tremelimumab or another anti-CTLA4 antibody [87, 88]. It is probable that this more careful strategy in trial design played a part in surpassing phase III and ultimately securing FDA approval for ipilimumab.

Conclusion

As we have seen, drug candidates can fail late in clinical development for a multitude of reasons. While some of these reasons may be difficult to predict, many that we have surveyed were due to foreseeable and thus avoidable errors, as summarized below:

- Foreseeable/avoidable errors: strategic and commercial factors; lack of sufficiently rigorous trial design; posthoc subgroup analysis identifying false biomarkers tested for

the first time in phase III; incomplete understanding of disease pathway; suboptimal dosing; proceeding directly from phase I to phase III trials; running multiple simultaneous clinical trials; improper inclusion criteria; inappropriate escape provisions; proceeding to phase III based on poor quality phase II data;

- Unforeseeable/unavoidable errors: target engagement that does not achieve disease modification; antidrug antibodies; preclinical results that do not translate to clinical outcomes.

What is perhaps somewhat discouraging is the fact that the late stage failures we reviewed were not isolated instances. Rather, the same errors were seen to repeat themselves in different development programs across the industry. It is perhaps suggestive of a degree of complacency and/or lack of urgency in addressing the present unsustainable rate of attrition in drug development. This may be because of the race to market, the fact that the costs of failed drug candidates are simply passed onto the public in the form of high drug prices on marketed products, or other reasons.

A limitation of the present study is that it was based solely on publicly available information. For many late stage failures, public disclosures are limited and therefore it is not possible to know what actually happened and learn from these failures.

Although the current rate of late stage failures in biologics development remains high, it is encouraging to note that many of the causes of such failures are avoidable and therefore there is much room for improvement. By learning from previous mistakes and adhering to the principles

and recommendations provided herein, it is possible to avoid these common pitfalls, increasing likelihood of success in phase III clinical trials and ultimately securing regulatory approval.

Acknowledgements

This work was carried out as Dr. Sun's research project within the Pharmaceutical Sciences pathway at UCSF.

Disclosure of Potential Conflicts of Interest

The authors report no conflict of interest.

Funding

Publication was funded in part by a Mary Anne-Koda-Kimble Seed Award for Innovation.

Author Contributions

We declare that this work was done by the authors named in this article and all liabilities pertaining to claims relating to the content of this article will be borne by the authors. A.S. and L.Z.B. designed the study. A.S. surveyed the literature. A.S. and L.Z.B. analyzed the information. A.S. and L.Z.B. wrote the manuscript.

Abbreviations

AD	Alzheimer's disease
ADAs	Antidrug antibodies
AE	Adverse event
AMD	Age-related macular degeneration
BAFF	B-cell activating factor
BLA	Biologics licensing application
CDMO	Contract development and manufacturing organization
CDR	Complementary-determining regions
CFI	Complement factor I
CI	Confidence interval
CSF	Cerebral spinal fluid
CTLA4	Cytotoxic T-lymphocyte antigen 4
DMARD	Disease-modifying anti-rheumatic drug
ESRD	End stage renal disease
FDAMA	Food and Drug Administration Modernization Act of 1997
FENO	Fraction exhaled nitric oxide
FFDCA	Federal Food, Drug Cosmetic Act of 1962
GA	Geographic atrophy
GI	Gastro-intestinal
LDL	Low-density lipoprotein

MACE	Major adverse cardiac events
MM	Multiple myeloma
MS	Multiple sclerosis
NSCLC	Non-small-cell lung cancer
OS	Overall survival
ORR	Overall response rate
PCSK9	Proprotein convertase subtilisin-kexin type 9
PD1	Programmed cell death 1
PET	Positron emission tomography
PFS	Progression-free survival
PK	Pharmacokinetic
PS	Phospholipid phosphatidylserine
RA	Rheumatoid arthritis
SAE	Serious adverse event
SLE	Systemic lupus erythematosus
TNF	Tumor necrosis factor

ORCID

Leslie Z. Benet http:// orcid.org/0000-0002-9678-2371

References

1. PhRMA. Modernizing drug discovery, development & approval. 2016 Mar 31 (last accessed January 20, 2018). <http://www.phrma.org/report/policy-solutions-modernizing-drug-discovery-development-and-approval>
2. Pharmaceutical Technology. Counting the cost of failure in drug development. 2017 June 19 (last accessed January 20, 2018). <https://www.pharmaceutical-technology.com/features/featurecounting-the-cost-of-failure-in-drug-development-5813046/>
3. Amplion. Fail early, fail fast – and increase likelihood-of-approval. 2017 Feb 14 (last accessed January 24, 2018). <http://www.amplion.com/biomarker-trends/fail-early-fail-fast-increase-likelihood-of-approval>
4. Seeking Alpha. Big pharma's mantra: fail early, fail fast. 2010 July 7 (last accessed January 24, 2018). <https://seekingalpha.com/article/213440-big-pharmas-mantra-fail-early-fail-fast>
5. Hwang TJ, Carpenter D, Lauffenburger JC, Wang B, Franklin JM, Kesselheim AS. Failure of investigational drugs in late-stage clinical development and publication of trial tesults. JAMA Internal Med. 2016;176(12):1826-1833.
6. Parasrampuria DA, Benet LZ, Sharma A. Why drugs fail in late stages of development: Case study analyses from the last decade and recommendations. AAPS J. 2018;20(3):46.
7. Paul J. Reducing the risk of late-stage failures (last accessed March 5, 2018). <https://pharma.elsevier.com/pharma-rd/reducing-the-risk-of-late-phase-failures/>

8. Merck and Pfizer's Phase III lung cancer trial of avelumab fails (last accessed March 5, 2018). <https://www.drugdevelopment-technology.com/news/merck-pfizers-phase-iii-lung-cancer-trial-avelumab-fails/>
9. Seattle Genetics discontinues Phase 3 CASCADE trial of vadastuximab talarine (SGN-CD33A) in frontline acute myeloid lymphoma (last accessed March 5, 2018). <https://www.businesswire.com/news/home/20170619005466/en/Seattle-Genetics-Discontinues-Phase-3-CASCADE-Trial>
10. Morrow T. Defining the difference: what makes biologics unique. Biotechnol Healthc. 2004;1:24-29.
11. BioWorld. Biologics' share of medicine chest grows, and so do the pricing concerns (last accessed January 24, 2018). <http://www.bioworld.com/content/biologics-share-medicine-chest-grows-and-so-do-pricing-concerns>
12. U.S. News & World Report. Biologics: the drugs transforming medicine. 2017 July 25 (last accessed January 24, 2018). <https://www.usnews.com/news/healthcare-of-tomorrow/articles/2017-07-25/biologics-the-drugs-that-are-transforming-medicine>
13. Rhoades W, Dickson D, Do DV. Potential role of lampalizumab for treatment of geographic atrophy. Clin Ophthalmol. 2015;9:1049-1056.
14. Holz FG, Sadda SR, Busbee B, Chew EY, Mitchell P, *et al*. Efficacy and safety of lampalizumab for geographic atrophy due to age-related macular degeneration. JAMA Ophthalmol. 2018;136(6):666-677.

15. Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Statistics in medicine – reporting of subgroup analyses in clinical trials. New Engl J Med. 2007;357(21):2189-2194.
16. Burke JF, Sussman JB, Kent DM, Hayward RA. Three simple rules to ensure reasonably credible subgroup analyses. BMJ 2015;351:h5651.
17. Rhoades W, Dickson D, Do DV. Potential role of lampalizumab for treatment of geographic atrophy [Retraction]. Clin Ophthalmology 2015;9:1135.
18. Spigel DR, Ervin TJ, Ramlau RA, Daniel DB, Goldschmidt JH, *et al*. Randomized phase II trial of onartuzumab in combination with erlotinib in patients with advanced non-small cell lung cancer. J Clin Oncol. 2013;31:4105-4114.
19. Iveson T, Donehower RC, Davidenko I, Tjulandin S, Deptala A, *et al*. Rilotumumab in combination with epirubicin, cisplatin, and capecitabine as first-line treatment for gastric or oesophagogastric junction adenocarcinoma: an open-label dose-descalation phase 1b study and a double-blind, randomized phase 2 study. Lancet Oncol. 2014;15:1007-1018.
20. Spigel DR, Edelman MJ, O'Byrne K, Paz-Ares L, Mocci S, *et al*. Results from the phase III trial of onartuzumab plus erlotinib versus erlotinib in previously treated stage IIIB or IV non-small-cell lung cancer: METLung. J Clin Oncol. 2016;35:412-420.
21. Shah MA, Bang YJ, Lordick F, Alsina M, Chen M, *et al*. Effect of fluorouracil, leucovorin, and oxaliplatin with or without onartuzumab in HER2-negative, MET-positive gastroesophageal adenocarcinoma: The METGastric randomized clinical trial. JAMA Oncol. 2017;3(5):620-627.
22. Catenacci DVT, Tebbutt NC, Davidenko I, Murad AM, Al-Batran SE, *et al*. Rilotumumab plus epirubicin, cisplatin, and capecitabine as first-line therapy in advanced

MET-positive gastric or gastro-oesophageal junction cancer (RILOMET-1): a randomised, double-blind, placebo-controlled phase 3 trial. Lance Oncol. 2017;18:1467-1482.

23. Ridker PM, Revkin J, Amarenco P, Brunell R, Curto M, *et al*. Cardiovascular efficacy and safety of bococizumab in high-risk patients. New Engl J Med. 2017;376:1527-1539.
24. Ballantyne CM, Neutel J, Cropp A, Duggan W, Wang EQ, *et al*. Results of bococizumab, a monoclonal antibody against proprotein convertase subtilisin/kexin type 9, from a randomized, placebo-controlled, dose-ranging study in statin-treated subjects with hypercholeseterolemia. Am J Cardiol. 2015;115:1212-1221.
25. Ridker PM, Revkin J, Amarenco P, Brunell R, Curto M, *et al*. Lipid-reduction variability and antidrug-antibody formation with bococizumab. New Engl J Med. 2017;376:1527-1539.
26. Pfizer discontinues global development of bococizumab, its investigational PCSK9 inhibitor. 2016 Nov 1. <https://www.pfizer.com/news/press-release/press-release-detail/pfizer_discontinues_global_development_of_bococizumab_its_investigational_pcsk9_inhibitor>
27. Pfizer abandons cholesterol drug bococizumab, writes it off as an earnings loss. 2016 Nov 3. <https://www.biospace.com/article/pfizer-abandons-cholesterol-drug-bococizumab-writes-it-off-as-an-earnings-loss-/>

28. Pfizer ends development of its PCSK9 inhibitor. 2016 Nov 1. <http://www.cardiobrief.org/2016/11/01/pfizers-ends-development-of-its-pcsk9-inhibitor/>

29. Harding FA, Stickler MM, Razo J, DuBridge RB. The immunogenicity of humanized and fully human monoclonal antibodies. mAbs. 2010;2(3):256-265.

30. Sheiner LB. Learning versus confirming in clinical drug development. Clin Pharmacol Ther. 1997;61(3):275-291.

31. Panettieri RA, Sjöbring U, Péterffy A, Wessman P, Bowen K. Tralokinumab for severe, uncontrolled asthma (STRATOS 1 and STRATOS 2): two randomised, double-blind, placebo-controlled, phase 3 clinical trials. Lancet Respir Med. 2018;6:511-525.

32. Isenberg DA, Petri M, Kalunian K, Tanaka Y, Urowitz MB, *et al*. Efficacy and safety of subcutaneous tabalumab in patients with systemic lupus erythematosus: results from ILLUMINATE-1, a 52-week, phase III, multicenter, randomised, double-blind, placebo-controlled study. Ann Rheum Dis. 2016;75:323-331.

33. Merrill JT, von Vollenhoven RF, Buyon JP, Furie RA, Stohl W, *et al*. Efficacy and Safety of subcutaneous tabalumab, a monoclonal antibody to B-cell activating factor, in patients with systemic lupus erythematosus: results from ILLUMINATE-2, a 52-week, phase III, multicenter, randomised, double-blind, placebo-controlled study. Ann Rheum Dis. 2016;75:332-340.

34. Schiff M, Combe B, Dörner T, Kremer JM, Huizinga TW. Efficacy and safety of tabalumab, an anti-BAFF monoclonal antibody, in patients with moderate-to-severe rheumatoid arthritis and inadequate response to TNF inhibitors: results of a randomised, double-blind, placebo-controlled phase 3 study. RMD Open. 2015;1:e000037.

35. Smolen JS, Weinblatt ME, van der Heijde D, Rigby WFC, van Vollenhoven R. Efficacy and safety of tabalumab, an anti-B-cell activating factor monoclonal antibody, in patients with rheumatoid arthritis who had an inadequate response to methotrexate therapy: results from a phase III multicentre, randomised, double-blind study. Ann Rheum Dis. 2015;74:1567-1570.

36. Lilly discontinues phase 3 rheumatoid arthritis program for tabalumab based on efficacy results. 2013 Feb 7. <https://investor.lilly.com/news-releases/news-release-details/lilly-discontinues-phase-3-rheumatoid-arthritis-program>

37. Lilly to discontinue development of tabalumab based on efficacy results in phase 3 lupus studies. 2014 Oct 2. <https://investor.lilly.com/news-releases/news-release-details/lilly-discontinue-development-tabalumab-based-efficacy-results>

38. Witcher J, Fleischmann R, Chindalore VL, Hansen RJ, Hu L, *et al*. Pharmacokinetics and safety of single doses of tabalumab in subjects with rheumatoid arthritis or systemic lupus erythematosus. Br J Clin Pharmacol. 2016;81:908-917.

39. Genovese MC, Bojin S, Biagini IM, Mociran E, Cristei D, *et al*. Tabalumab in rheumatoid arthritis patients with an inadequate response to methotrexate and naïve to biologic therapy. Arthritis Rheum. 2013;65(4):880-889.

40. Genovese MC, Fleischmann RM, Greenwald M, Satterwhite J, Veenhuizen M. Tabalumab, an anti-BAFF monoclonal antibody, in patients with active rheumatoid arthritis with an inadequate response to TNF inhibitors.

41. Genovese MC, Lee E, Satterwhite J, Veenhuizen M, Disch D, *et al*. A phase 2 dose-ranging study of subcutaneous tabalumab for the treatment of patients with active rheumatoid arthritis and an inadequate response to methotrexate. Ann Rheum Dis. 2013;72:1453-1460.
42. Greenwald M, Szczepanski L, Kennedy A, Veenhuizen M, Komocsar WJ, *et al*. A 52-week, open-label study evaluating the safety and efficacy of tabalumab, an anti-B-cell activating factor monoclonal antibody, for rheumatoid arthritis. Arthritis Res Ther. 2014;16:415.
43. Mujtaba MA, Komocsar WJ, Nantz E, Samaniego MD, Henson SL, *et al*. Effect of treatment with tabalumab, a B-cell activating factor inhibitor, on highly sensitized patients with end-stage renal disease awaiting transplantation. Am J Transplant. 2016;16:1266-1275.
44. GSK announces regulatory submissions for subcutaneous formulation of Benlysta® (belimumab) for patients with systemic lupus disease. 2016 Sep 23. <https://us.gsk.com/en-us/media/press-releases/2016/gsk-announces-regulatory-submissions-for-subcutaneous-formulation-of-benlysta-belimumab-for-patients-with-systemic-lupus-disease/>
45. Cummings JL, Morstorf T, Zhong K. Alzheimer's disease drug pipeline: few candidates, frequent failures. Alzheimers Res Ther. 2014;6:37.
46. DeMattos RB, Bales KR, Cummins DJ, Dodart JC, Paul SM, *et al*. Peripheral anti-Aβ antibody alters CNS and plasma Aβ clearance and decreases brain Aβ burden in a mouse model of Alzheimer's disease. PNAS. 2001;98(15):8850-8855.

47. Elder GA, Gama Sosa MA, De Gasperi R. Transgenic mouse models of Alzheimer's Disease. Mt Sinai J Med. 2010;77(1):69-81.

48. Siemers ER, Friedrich S, Dean RA, Gonzales CR, Farlow MR, *et al*. Safety and changes in plasma and cerebrospinal fluid amyloid β after a single administration of an amyloid β monoclonal antibody in subjects with Alzheimer Disease. Clin Neuropharmacol. 2010;33(2):67-73.

49. Farlow M, Arnold SE, van Dyck CH, Aisen PS, Snider BJ, *et al*. Safety and biomarker effects of solanezumab in patients with Alzheimer's Disease. Alzheimers Dement. 2012;8:261-271.

50. Doody RS, Thomas RG, Farlow M, Iwatsubo T, Vellas B. Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's Disease. New Engl J Med. 2014;370:311-321.

51. Honig LS, Vellas B, Woodward M, Boada M, Bullock R, *et al*. Trial of solanezumab for mild dementia due to Alzheimer's Disease. New Engl J Med. 2018;378:321-330.

52. Siemers ER, Sundell KL, Carlson C, Case M, Sethuraman G, *et al*. Phase 3 solanezumab trials: Secondary outcomes in mild Alzheimer's disease patients. Alzheimers Dement. 2016;12:110-120.

53. Ioannidis JPA. Extrapolating from Animals to Humans. Sci Transl Med. 2012;4:151ps15.

54. Has Alzheimer's research been wrong for 20 years? 2017 May 15. <https://www.alzheimers.net/2014-04-07/beta-amyloid-wrong-protein/>

55. Is the leading theory about Alzheimer's wrong? 2017 Feb 22. <https://www.theatlantic.com/health/archive/2017/02/alzheimers-amyloid-hypothesis/517185/>

56. There's new evidence that the dominant approach to Alzheimer's may be completely wrong. 2018 Aug 13. <https://www.marketwatch.com/story/theres-new-evidence-that-the-dominant-approach-to-alzheimers-may-be-completely-wrong-2018-08-09>

57. Corren J, Lemanske RF, Hanania NA, Korenblat PE, Parsey MV, *et al*. Lebrikizumab treatment in adults with asthma. N Engl J Med. 2011;365:1088-1098.

58. Hanania NA, Korenblat P, Chapman KR, Bateman ED, Kopecky P, *et al*. Efficacy and safety of lebrikizumab in patients with uncontrolled asthma (LAVOLTA I and LAVOLTA II): replicate, phase 3, randomised, double-blind, placebo-controlled trials. Lancet Resp Med. 2016;4:781-796.

59. Wenzel S, Castro M, Corren J, Maspero J, Wang L, *et al*. Dupilumab efficacy and safety in adults with uncontrolled persistent asthma despite use of medium-to-high dose inhaled corticosteroids plus a long-acting β2 agonist: a randomised, double-blind placebo-controlled pivotal phase 2b dose-ranging trial. Lancet 2016;388:31-44.

60. Castro M, Corren J, Pavord ID, Maspero J, Wenzel S, *et al*. Dupilumab efficacy and safety in moderate-to-severe uncontrolled asthma. N J Engl Med. 2018;378:2486-2496.

61. FDA approves asthma indication for Dupixent® (dupilumab). 2018 Oct 19 (last accessed February 24, 2019). <https://newsroom.regeneron.com/news-releases/news-release-details/fda-approves-asthma-indication-dupixentr-dupilumab>

62. Gerber DE, Horn L, Boyer M, Sanborn R, Natale R, *et al*. Randomized phase III study of docetaxel plus bavituximab in previously treated advanced non-squamous non-small-cell lung cancer. Ann Oncol. 2018;29:1548-1553.

63. Gerber DE, Spigel DR, Giorgadze D, Shtivelband M, Ponomarova OV, *et al*. Docetaxel combined with bavituximab in previously treated, advanced nonsquamous non-small-cell lung cancer. Clin Lung Cancer. 2016;17(3):169-176.

64. Peregrine Pharmaceuticals, Inc. Corporate Fact Sheet. 2014 Jan (last accessed December 4, 2018). <http://files.shareholder.com/downloads/PPHM/0x0x621302/6cfcdcd0-7d2d-4395-89cb-0d33fc4a8c85/jan_corp_fact_sheet.pdf>

65. Laforest R, Dehdashti F, Liu Y, Frye J, Frye S, *et al*. First-in-man evaluation of ^{124}I-PGN650: A PET tracer for detecting phosphatidylserine as a biomarker of the solid tumor microenvironment. Mol Imaging 2017;16:1-9.

66. Williams S. Just how important is bavituximab to Peregrine Pharmaceuticals? 2014 Mar 11 (last accessed December 4, 2018) <https://www.fool.com/investing/general/2014/03/11/just-how-important-is-bavituximab-to-peregrine-pha.aspx>

67. Feuerstein A. Peregrine Pharma stock craters on lung cancer drug blowup. 2012 Sep 24 (last accessed December 4, 2018) <https://www.thestreet.com/story/11716746/1/peregrine-pharma-stock-crater-on-lung-cancer-drug-blowup.html>

68. Avid offloads phase III mAb, becomes pure play CDMO. 2018 Feb 20 (last accessed December 4, 2018) <https://www.biopharma-reporter.com/Article/2018/02/20/Avid-offloads-Phase-III-mAb-becomes-pureplay-CDMO>

69. Peregrine Pharmaceuticals announces name change to Avid Bioservices as part of transition to dedicated contract development and manufacturing organization (CDMO). 2018 Jan 5 (last accessed December 4, 2018) <http://ir.avidbio.com/news-releases/news-release-details/peregrine-pharmaceuticals-announces-name-change-avid-bioservices?ReleaseID=1053451>

70. Takeuchi T, Thorne C, Karpouzas G, Sheng S, Xu W, *et al*. Sirukumab for rheumatoid arthritis: the phase III SIRROUND-D study. Ann Rheum Dis 2017;76:2001-2008.

71. Aletaha D, Bingham CO, Tanaka Y, Agarwal P, Kurrasch R, *et al.* Efficacy and safety of sirukumab in patients with active rheumatoid arthritis refractory to anti-TNF therapy (SIRROUND-T): a randomized, double-blind, placebo-controlled, parallel-group, multinational, phase 3 study. Lancet 2017;389:1206-1217.

72. Taylor PC, Schiff MH, Wang Q, Jiang Y, Zhuang Y, *et al.* Efficacy and safety of monotherapy with sirukumab compared with adalimumab monotherapy in biologic-naïve patients with active rheumatoid arthritis (SIRROUND-H): a randomized, double-blind, parallel-group, multinational, 52-week, phase 3 study. Ann Rheum Dis 2018;77:658-666.

73. Janssen submits application seeking approval of sirukumab in United States for rheumatoid arthritis. 2016 Sep 23 (last accessed December 5, 2018). <https://www.jnj.com/media-center/press-releases/janssen-submits-application-seeking-approval-of-sirukumab-in-united-states-for-rheumatoid-arthritis>

74. Worcester S. FDA committee rejects sirukumab approval on safety concerns. 2017 Aug 2 (last accessed December 5, 2018). <https://www.mdedge.com/rheumatologynews/article/143848/rheumatoid-arthritis/fda-committee-rejects-sirukumab-approval-safety>

75. FDA Arthritis Advisory Committee (AAC) FDA Introductory Remarks, BLA 761057: Sirukumab for the treatment of adult patients with moderately to severely active rheumatoid arthritis (RA) who have had an inadequate response or are intolerant to one or more disease modifying anti-rheumatic drugs (DMARDs). 2017 Aug 2 (last accessed December 6, 2018). <https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/UCM570356.pdf>

76. Janssen receives complete response letter from U.S. FDA for sirukumab biologics license application. 2017 Sep 22 (last accessed December 5, 2018). <https://www.jnj.com/media-center/press-releases/janssen-receives-complete-response-letter-from-us-fda-for-sirukumab-biologics-license-application>

77. J&J halts sirukumab development, talacotuzumab trial. 2017 Oct 17 (last accessed December 5, 2018). <https://www.genengnews.com/topics/translational-medicine/jj-halts-sirukumab-development-talacotuzumab-trial/>

78. FDA Guidance for Industry, Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products. 1998 May (last accessed December 9, 2018). <https://www.fda.gov/ucm/groups/fdagov-public/@fdagov-drugs-gen/documents/document/ucm072008.pdf>

79. Sirukumab presentation to the Arthritis Advisory Committee, Janssen R&D, LLC. 2017 Aug 2 (last accessed December 8, 2018). <https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/UCM570357.pdf>

80. Summary minutes of the Arthritis Advisory Committee Meeting. 2017 Aug 2 (last accessed December 9, 2018). <https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/UCM575678.pdf>

81. FDA Briefing Document, Arthritis Advisory Committee Meeting. 2017 Aug 2 (last accessed December 8, 2018). <https://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/ArthritisAdvisoryCommittee/UCM569150.pdf>

82. Smolen JS, Aletaha D, McInnes IB. Rheumatoid Arthritis. Lancet 2016;388:2023-2038.

83. Ma X, Xu S. TNF inhibitor therapy for rheumatoid arthritis (review). Biomed Rep. 2018;1:177-184.

84. Singh JA, Saag KG, Bridges SL, Akl EA, Bannuru RR, *et al*. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. Arthritis Rheumatol. 2016;68(1):1-26.

85. AstraZeneca provides update on the phase III MYSTIC trial of Imfinzi and tremelimumab in stage IV non-small cell lung cancer. 2018 Nov 16 (last accessed February 24, 2019) <https://www.astrazeneca.com/media-centre/press-releases/2018/astrazeneca-provides-update-on-the-phase-iii-mystic-trial-of-imfinzi-and-tremelimumab-in-stage-iv-non-small-cell-lung-cancer16112018.html>

86. Taylor NP. AstraZeneca's tremelimumab fails another phase 3 cancer trial. 2018 Dec 7 (last accessed February 24, 2019) <https://www.fiercebiotech.com/biotech/astrazeneca-s-tremelimumab-fails-another-phase-3-cancer-trial>

87. Ribas A, Kefford R, Marshall MA, Punt CJA, Haanen JB, *et al*. Phase III randomized clinical trial comparing tremelimumab with standard-of-care chemotherapy in patients with advanced melanoma. J Clin Oncol. 2013;31:616-622.

88. Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, *et al*. Improved survival with ipilimumab in patients with metastatic melanoma. N Engl J Med. 2010;363:711-723.

Table 1. Recent failures of large molecule mAbs drugs in phase III clinical trials (2014-2019)

Drug	Sponsor	Indication
Aducanumab	Biogen	Alzheimer's disease
Anifrolumab	AstraZeneca	Systemic lupus erythematosus
Bavituxumab	Peregrine	Non-small cell lung cancer
Bimagrumab	Novartis	Sporadic inclusion body myositis
Bococizumab	Pfizer	Hyperlipidemia
Crenezumab	Genentech	Alzheimer's disease
Fulranumab	Johnson & Johnson	Osteoarthritis pain
Lampalizumab	Genentech	Geographic atrophy
Lebrikizumab	Genentech	Asthma
Mirvetuximab	ImmunoGen	Ovarian cancer
NEOD001	Prothena	Amyloidosis
Onartuzumab	Roche	Non-small cell lung cancer; gastroesophageal cancer
Rilotumumab	Amgen	Gastric and gastroesophageal cancer
Sirukumab	Janssen	Rheumatoid arthritis
Solanezumab	Eli Lilly	Alzheimer's disease
Suptavumab	Regeneron	Respiratory synctitial virus
Tabalumab	Eli Lilly	Systemic lupus erythematosus; rheumatoid arthritis
Talacotuzumab	Johnson & Johnson	Acute myeloid leukemia
Tralokinumab	AstraZeneca	Asthma

Tremelimumab	AstraZeneca	Metastatic melanoma; head and neck squamous cell carcinoma; non-small cell lung cancer
Vadastuximab	Seattle Genetics	Acute myeloid leukemia

Table 2. Tabalumab Clinical Development Program

Indication	Phase I	Phase II	Phase III
RA	NCT01253226 NCT01253291 H9B-LC-BCDB H9B-MC-BCDE	NCT00308282 NCT00689728 NCT00785928 NCT00837811 NCT01576549	NCT01198002 NCT01202760 NCT01202773 NCT01215942 NCT01676701
SLE	H9B-LC-BCDB		NCT01196091 NCT01205438 NCT01488708 NCT02041091
MM	NCT00689507 NCT01556438	NCT01602224	
MS		NCT00882999	
ESRD		NCT01200290	

RA = rheumatoid arthritis, SLE = systemic lupus erythematosus, MM = multiple myeloma, MS = multiple sclerosis, ESRD = end stage renal disease.