Japanese Journal of Gastroenterology and Hepatology

Review Article

ISSN 2435-1210 |Volume 7

Complex Innovative Design for Liver Disease Clinical Trials

Chow SC[1], Mai W[1*], Suzuki A[1], Kong Y[2] and Jia J[2]

[1]Duke University School of Medicine, Durham, North Carolina, USA

[2]Beijing Friendship Hospital, Capital Medical University, Beijing, China

[*]Corresponding author:

Weijia Mai,
Department of Biostatistics and Bioinformatics, Duke University School of Medicine, 2424 Erwin Road, Durham, North Carolina, USA,
E-mail: weijia.mai@duke.edu;
sheinchung.chow@duke.edu

Received: 05 Sep 2021
Accepted: 29 Sep 2021
Published: 05 Oct 2021

Keywords:

Endpoint selection; Population shift; Therapeutic index; Seamless adaptive design

Citation:

Mai W,, Complex Innovative Design for Liver Disease Clinical Trials. Clinical Case Report. Japanese J Gstro Hepato. 2021; V7(4): 1-13

1. Abstract

In recent years, the prevalence of liver diseases such as hepatitis B (or C) virus infection (HBV or HCV), non-alcoholic fatty liver disease (NAFLD) and non-alcoholic steatohepatitis (NASH) have received much attention world-wide. Due to the high prevalence and the health burden, there is an unmet medical need to develop therapeutic strategies for patients with potentially progressing liver diseases such as NASH. For development of treatment for NASH, the primary objective is to prevent liver-related morbidity and mortality, which generally takes a long time (e.g., more than 10-20 years) to develop. In practice, it is not feasible to conduct such clinical studies. Thus, regulatory agencies such as United States Food and Drug Administration (FDA) suggests the use of complex innovative design to (i) provide an unbiased and reliable assessment of the treatment effect of the test drug under investigation, (ii) shorten the development process, and (iii) speed up the overall path to regulatory review and approval. In this article, several complex innovative designs including a two-stage seamless adaptive trial design are proposed. Relative advantages and limitations of these complex innovative designs are discussed from both regulatory and clinical/statistical perspectives. Critical statistical/clinical issues and a couple of case studies related to drug development of liver diseases such as HBV/HCV and NASH are also described and discussed.

2. Introduction

In recent years, the prevalence of liver-related morbidity and mortality caused by certain liver diseases is increasing world-wide [34, 42]. Thus, the development of treatment for liver diseases such as hepatitis B (or C) virus infection (HBV or HCV), non-alcoholic fatty liver disease (NAFLD) and non-alcoholic steatohepatitis (NASH) have received much attention. HBV and HCV infections are the most common causes of liver disease worldwide. Both viruses can be transmitted parenterally, sexually and perinatally, with perinatal and sexual transmission being more common for HBV than for HCV. More than 500 million people worldwide are persistently infected with the HBV and/or HCV and are at risk of developing cirrhosis and hepatocellular carcinoma [25, 33]; NAFLD is the presence of hepatic accumulation of triglycerides in the hepatocytes in the absence of significant alcohol intake, viral infection, or any other specific etiology of liver disease. NASH is defined histologically by the presence of hepatic steatosis with evidence for hepatocyte damage with or without fibrosis. The most important histological feature associated with mortality in NASH is presence of significant fibrosis. Therefore, NAFLD represents a histopathologic spectrum ranging from simple steatosis to NASH, fibrosis, and cirrhosis [34, 35]. Although recent data suggest that some patients with simple fatty liver can progress to NASH and clinically significant fibrosis, most of the fibrosis progression seems to occur in patients with NASH. NASH has been recognized as one of the leading causes of cirrhosis in adults and NASH related cirrhosis is currently the second indication for liver transplants in the United States [41].

Due to the increasing prevalence of NAFLD/NASH there is an unmet medical need to develop therapeutic interventions for the patients with such liver diseases. However, pharmacologic therapeutic interventions in these liver diseases (e.g., NASH) have largely proven

to be ineffective or unappealing due to long-term side-effect profiles, and the majority of patients cannot achieve targeted therapeutic effect [22]. Thus, the development and acceptance of meaningful, readily obtainable, and well-defined clinical trial endpoints are imperative to develop new and effective therapies to treat this growing epidemic. In liver disease clinical trials, however, study endpoints may be different depending upon disease status/stage and complications. For example, for NASH clinical trials, composite endpoint of (i) complete resolution of steatohepatitis and no worsening of fibrosis, or (ii) at least one point improvement in fibrosis with no worsening of steatohepatitis (no increase in steatosis, ballooning or inflammation) may be considered. In this article, without loss of generity and for illustration purpose, we will focus on design and analysis of NASH clinical trials.

While the majority of patients with simple steatosis generally have a benign course, a diagnosis of NASH carries a significantly higher risk for disease progression, cirrhosis, HCC and death. The complex molecular mechanisms leading to NASH and the long duration of time to develop complications of disease are challenges to developing meaningful clinical endpoints. Because of these challenges, surrogate endpoints that are linked to all-cause mortality, liver-related death, and complications of cirrhosis are much more likely to be beneficial in the majority of patients. The diagnosis of NASH requires a liver biopsy with subsequent confirmation of a specific histopathologic pattern. The minimal criteria for the diagnosis of steatohepatitis include the presence of >5 % macro vesicular steatosis, inflammation, and hepatocyte ballooning, typically with a predominantly centrilobular distribution in adults. The NAFLD activity score (NAS) is a validated scoring system that can be used to assess histologic change in studies of both adults and children with NASH [24]. It is an un-weighted composite of steatosis, inflammation, and ballooning scores. It should be noted that a score of 5 or more is associated with a greater likelihood of having NASH, but an NAS ≥5 does not necessarily confirm a diagnosis of NASH. Additionally, NAS has not been validated as a marker for likelihood of disease progression (e.g., cirrhosis, mortality) and/or response to therapy [3]. The use of NAS is currently limited to clinical trial settings. Thus, it is recommended that a validated method for the staging of NASH be used for assessment of changes in disease stage in clinical trials of NASH. In practice, the NASH Clinical Research Network (CRN) fibrosis staging system (ranging from 0 to 4 for no fibrosis to cirrhosis) is the most validated score system currently available.

Given the high unmet medical need, no universally accepted clinical endpoints, the lack of validated surrogate endpoints, and no approved drug therapy for NASH, the United States Food and Drug Administration (FDA) published guidance to assist the sponsor in drug development of NASH (FDA, 2018b). FDA's guidance provided some general considerations including (i) specify criteria that establish a diagnosis of cirrhosis, e.g., a diagnosis of cirrhosis should be supported by histology such as a NASH Clinical Research Network (CRN) fibrosis score of 4, (ii) stratified randomization, e.g., patients with type 2 diabetes mellitus or patients with NASH-cirrhosis who are treated with Vitamin E or pioglitazone, (iii) sufficient duration and adequate sample size, and (iv) establishment of expert committee to adjudicate cases for safety. In addition, FDA encourages the use of biochemical or imaging non-invasive biomarkers that can replace liver biopsies.

A few aspects unmentioned in the FDA guideline deserve attention. NAFLD occurs in both genders and spans a wide range of age groups. However, the risk and, probably, disease-driving mechanisms differ depending on age, sex, and reproductive status. As extensively reviewed in recent papers [27, 28], most of the key mechanisms involved in the NAFLD pathobiology exhibit sexual dimorphism. The liver is metabolically distinct between males and females, and some pathways commonly targeted in NAFLD drug development, such as FXR and PPARα, are differently regulated in males and females [13]. Thus, the consideration of age, sex, and reproductive status in the study design of clinical trials is crucial to assess potential disparities in drug efficacy and safety profile (e.g., stratification, blocked enrolment, and sufficient data collection) in drug development. Given the accumulating knowledge of sexual dimorphism in NAFLD and the fact that most preclinical studies for the drug development are still performed using male animals, not both sexes, the consideration of sex and reproductive status in a trial design and analysis should be institutionalized in future drug development while translating the animal findings into humans. In analysing trial data, sex and age are frequently assumed as biologically independent. However, biophysiological profiles in women are significantly altered after their age at menopause. Thus, a simple comparison between men and women without considering menopausal status is an insufficient analysis from a physiological viewpoint. Similarly, including age and sex in a model to adjust for a confounding effect without checking potential age-/sex-interaction is also inadequate and could generate misleading results.

As a result, it is suggested that a complex innovative design should be used to (i) provide an unbiased and reliable assessment of the treatment effect of the test drug under investigation, (ii) shorten the development process, and (iii) speed up the overall path to regulatory review and approval, while properly considering relevant biological disparities by age and sex. Along this line, it is suggested that a seamless adaptive design be used in NASH clinical studies [5, 9]. A seamless adaptive design provides the flexibility and efficiency for identifying potential signals of clinical benefit of the test treatment under investigation and make prompt pre-planned adaptations without undermining the validity or integrity of the trial [10, 20]. The adaptive design can also be applied to adequately consider potential sex-/age-disparities in a study design without negatively affecting the study cost and duration. For instance, a trial can employ randomization blocked by age and sex to assess the efficacy and safety between individual blocks comparatively; if no effects or poor risk/benefit in

a particular block, the block can be dropped from further enrolment. In NASH clinical trials, however, some practical issues inevitably occur. These issues include, but are not limited to, (i) only limited number of patients willing to have multiple liver biopsies, (ii) lack of validated surrogate endpoints, (iii) long-term exposure to assess an impact in outcomes. [21] indicated the use of seamless adaptive design appear reasonable for addressing these practical issues. A continuous seamless adaptive design may reduce the overall sample size while allowing some patients to continue after each one of the phases (if a multiple-stage seamless adaptive design is used).

In the next section, limitations and challenges in chronic liver disease clinical trials are described. Section 3 provide some general considerations in liver disease drug development from regulatory and clinical/statistical perspectives. Section 4 introduces complex innovative design such as a two-stage seamless adaptive trial design and its application in liver disease clinical studies. Section 5 provides a couple of case studies concerning liver disease clinical trials utilizing a two-stage seamless adaptive trial design. Some concluding remarks are given in the last section of this article.

3. Limitations and Challenges in Chronic Liver Disease Clinical Trials

Due to the complexity of the disease course, there are several limitations in the conduct of liver disease clinical trials. These limitations include the lack of accurate, reproducible, and easily applied methods to assess the test treatment under investigation, which creates major challenges not only for drug development, but also in the clinical management of patients with liver diseases. Without loss of generality and for illustration purpose, we will focus on clinical development for treating patients with NASH.

3.1. Disease Assessment

Liver biopsy remains the gold standard for the diagnosis of chronic liver diseases such as NASH but it has several limitations. There is always a risk that the biopsy taken might not be representative for the amount of fibrosis in the whole liver. Increasing the length of liver biopsy specimen decreases the risk of sampling error. In general, 15 mm is considered sufficient. Not only the length but also the caliber of the biopsy needle is important in order to obtain a piece of liver of adequate size for histological evaluation: a 16-gauge needle is considered appropriate for percutaneous liver biopsy [35]. Inter-observer variation is another limitation which is related to the discordance between pathologists in biopsy interpretation. This can be as high as 25% but the variation is less pronounced when the biopsy assessment is done by specialized liver pathologists (Papastergiou, Tsochatzis, and Burroughs, 2012).

3.2. Study Objectives and Endpoints

For HCV/HBV clinical studies, sustained virologic response (SVR) at 72 weeks (i.e., 24 weeks follow-up after 48 weeks of treatment) is considered acceptable efficacy endpoint for HCV clinical trials, while the following efficacy endpoints are suggested for HBV clinical trials (FDA, 2018a): (i) suppression of HBV DNA on-treatment, (ii) sustained suppression (more than 6 months) of HBV DNA off treatment after a finite duration of therapy, and (iii) sustained suppression (more than 6 months) of HBV DNA off treatment with HBsAg loss (less than 0.05 international unit/milliliter (IU/mL)) with or without HBsAb seroconversion after a finite duration of therapy.

For NASH clinical development, on the other hand, composite endpoint of (i) complete resolution of steatohepatitis and no worsening of fibrosis, or (ii) at least one point improvement in fibrosis with no worsening of steatohepatitis (no increase in steatosis, ballooning or inflammation) may be considered. The primary objective of the treatment for NASH is to prevent liver-related morbidity and mortality, due mainly to the development of cirrhosis, which generally takes more than 10-20 years to develop. Due to this long natural history, there is a need of surrogate markers of avoidance of cirrhosis and thus liver-related mortality. The main predictor of disease progression is increasing fibrosis. Therefore, patients with steatohepatitis are more likely to have a progressive disease than patients with simple fatty liver. Therefore, complete resolution of NASH (i.e. absence of ballooning with no or minimal inflammation) with no worsening of fibrosis, or actual improvement in fibrosis are recommended as "surrogate endpoints, reasonably likely to predict clinical benefit in progression to cirrhosis and liver-related death" [36] (Table 1). Ideally, a co-primary endpoint of two composite endpoints, 1-complete resolution of NASH with no worsening of fibrosis and, 2-at least one-point improvement in the fibrosis stage with no worsening of steatohepatitis, should be demonstrated in the clinical trials for NDA. However, the clinical outcome study, aiming to demonstrate a reduction in progression to cirrhosis and portal hypertension/cirrhosis related events needs to be demonstrated after marketing authorization.

For dose ranging trials, the histology endpoint of improvement in activity as assessed by a reduction in at least 2 points in NAS (including at least 1 point in ballooning or inflammation) is an acceptable surrogate marker of improvement. It is important to note that though NAS has proven useful for comparative analyses and interventional studies it does not provide information about fibrosis or the location of lesions. Therefore, the reduction in *NAS must be associated with lack of progression in fibrosis.*

Biopsy-based endpoints are, in general, not feasible in a 12-24-week POC trials. In short-term POC studies, designed mainly to assess tolerability of new drugs and to look for futility signals to direct decisions regarding further development, an improvement of hepatic steatosis, as determined by magnetic resonance technology, might be suitable since improvement in steatohepatitis is generally associated with a reduction in liver fat [35]. Improvement in liver chemistry (e.g., aminotransferase) and other non-invasive biomarkers of insulin sensitivity, inflammation, apoptosis and fibrosis could be helpful to evaluate the efficacy of the compound and support decision making. However, it is important to note that the use of non-invasive bio-

markers methods is still considered experimental and there are no validated non-invasive biomarkers.

3.3. Target Population

Prevention of cirrhosis and demonstrating a positive effect on well-defined liver outcomes are key clinical goals when considering a NASH drug development program. Therefore, for trials aiming to support marketing application, it is important that subjects with the greatest risk of progression to cirrhosis are enrolled (Table 1). Among individual features, liver fibrosis has proven the best independent association with liver-related as well as overall mortality. Patients with NASH develop progressive fibrosis in 25% to 50% of individuals over 4-6 years, while 15% to 25% of individuals with NASH can progress to cirrhosis [30]. In another study, with a mean follow-up of 13 years, 13.3% of NASH patients with mild to moderate fibrosis (stage 1-2) and 50% of patients with fibrosis stage 3 at inclusion developed cirrhosis [15]. Since in patients with NASH and fibrosis stage F2-F3, the probability of developing cirrhosis is much higher than in patients with early fibrosis (F1), this population is recommended for long-term outcome trials in order to enhance the chances of demonstrating a benefit within a reasonable timeframe. Furthermore, age, sex, and menopausal status affect the susceptibility of the fibrotic response [23, 27]. Pre-menopausal women with sufficient physiological oestrogens are protected from fibrogenesis, even when they have active NASH features. Thus, premenopausal women may not be good candidates for a long-term outcome trial assessing the effect of an agent on the fibrosis progression, unless accompanied by a high-risk condition, such as cystic ovary syndrome. Again, since baseline risk of fibrosis are different by age, sex, and menopausal status, participants with these factors should be distributed equally between the arms. The enrolment of patients with moderate /advanced fibrosis for the evaluation of long-term outcomes including progression to cirrhosis should ensure that an expected number of events, calculated based on progression rate for each fibrosis stage, are obtained based on the literature [1, 15, 2, 31, 38] in patients with NASH and advanced fibrosis (F2-F3) this progression rate can be estimated at 8% per year for fibrosis stage 3, and 6% per year for fibrosis stage 2. Since the progression rate in some patients with mild fibrosis with additional risk factors of progression (e.g. presence of type 2 diabetes, the metabolic syndrome, high transaminases) might be fast, it is worth exploring this subgroup of patients, as an additional exploratory group.

A broad population of NASH patients including those ones with mild fibrosis is acceptable in dose ranging (phase 2) trials. However, it is recommended that a sufficient number of patients with moderate and severe fibrosis are enrolled in order to get preliminary data to inform the trial/s to support marketing application. Ideally, in early proof of concept trials, the target population should also be patients with biopsy confirmed NASH. However, patients at high risk of NASH, namely patients with fatty liver and diabetes and/or the metabolic syndrome with or without high liver enzymes can be acceptable at this stage. Non-invasive serum biomarkers or imaging can be used to enrich a population in a POC trial.

4. General Considerations

As discussed above, there is unmet medical need for treatment of critical liver diseases. In addition, due to the complexity of the liver diseases such as NASH, there exists no approved drugs in the United States. Thus, to ensure the success of clinical development for treating critical liver diseases, it is suggested that some strategic frameworks should be implemented. These strategic frameworks include potential use of surrogate or biomarker endpoints as well as innovative statistical methods for assessment of the test treatment under investigation. These logistical approaches could be considered for applying complex innovative designs to clinical trials with the goal of facilitating drug development for this growing unmet medical need. In this section, general considerations for development of treatment for liver diseases are discussed from both regulatory and clinical/statistical perspectives.

4.1 Regulatory Perspective and Guidance

In 2018, FDA published a couple of guidance to assist the sponsors in development of drugs for treatment of HBV and NASH (FDA, 2018a, 2018b). In practice, given the high prevalence of these liver diseases, the associated morbidity, the growing burden of end-stage liver disease, and the limited availability of livers for organ transplantation, the 2018 FDA draft guidance provided recommendations for pre-clinical and clinical development including trial design and endpoint selection to support regulatory approval of drugs to treat patients with HBV/HCV [18] or NAFLD/NASH [19]. Without loss of generality and for illustration purpose, as an example, we will focus on the FDA 2018 draft guidance on development of drug products for treating NASH. This guidance is meant to assist the sponsor in identifying therapies that will slow the progress, halt, or reverse NASH and NAFLD, which will address an unmet medical need [19]. However, FDA indicated that the guidance [19] is not meant to cover the development of drugs to treat cirrhosis caused by NASH or the development of *in vitro* diagnostics that may be used in developing drugs to treat the disease. Some basic considerations are briefly described below.

4.2. Basic Considerations

As stated in the draft guidance [19], NAFLD consists of three successive stages: non-alcoholic fatty liver (NAFL), noncirrhotic NASH, and NASH with cirrhosis. The draft guidance provides sponsors a convenient conceptual framework to identify areas of potential future drug development. However, because patients' NAFL can exist for many years and may not progress to NASH, it may be challenging to demonstrate a favourable benefit-risk profile of pharmacological treatment(s) in NAFL patients. Thus, it is suggested that sponsors should consider the following general considerations during drug development for treatment of noncirrhotic NASH with liver fibrosis [19]:

First, the sponsor should consider using animal models for NASH to screen and identify potential investigational drugs. The sponsor should select a specific animal model based on the mechanism of action of the investigational drug. Ideally, both males and females should be assessed in preclinical experiments from the above-mentioned reasons to inform proper consideration of sex and reproductive status in a trial for the efficacy and safety assessment. Second, if there is a potential for liver toxicity based on animal toxicology studies, the sponsor should institute an appropriate plan to monitor liver safety early in drug development and to collect sufficient data/information allowing accurate causality assessment if any events occur. For such a plan, the sponsor should consider the challenges of effectively recognizing a liver signal in a chronic liver condition such as NASH, collecting more data points of baseline liver chemistry data.

As stated in the guidance, until a sponsor can characterize a drug's initial tolerability, preliminary safety, and pharmacokinetics, patients with evidence of abnormal liver synthetic function should be excluded from early phase trials (i.e., phase 1 and early proof-of-concept (POC) clinical trials). In addition, the sponsor should study the effects of hepatic impairment on the drug's pharmacokinetics early during the drug development program in a dedicated hepatic study to support appropriate dosing and dose adjustment across the spectrum of NASH liver disease.

4.3. Specific Considerations

In addition to basic considerations, the guidance also provided specific considerations regarding phase 2 development considerations, phase 3 development considerations and pediatric considerations, which are summarized below.

For phase 2 and phase 3 studies, FDA suggested that sponsors should enrol patients with a histological diagnosis of NASH with liver fibrosis made within six months of enrolment, taking into consideration patients' standard of care and background therapy for other chronic conditions. FDA also says that patients' weight should be stable for three months prior to enrolment. Furthermore, FDA indicated that phase 3 studies for NASH should be double-blind and placebo-controlled with the goal of slowing, halting or reversing disease progression and improving clinical outcomes.

Because of the slow progression of NASH and the time required to conduct a trial that would evaluate clinical endpoints such as progression to cirrhosis or survival, the FDA recommends sponsors consider several liver histological improvements as endpoints reasonably likely to predict clinical benefit to support accelerated approval under the regulations. FDA recommended endpoints include (i) resolution of steatohepatitis on overall histopathological reading and no worsening of liver fibrosis on NASH CRN fibrosis score, (ii) improvement in liver fibrosis greater than or equal to one stage (NASH CRN fibrosis score) and no worsening of steatohepatitis (defined as no increase in NAS for ballooning, inflammation, or steatosis), or (iii) both resolution of steatohepatitis and improvement in fibrosis. Although the liver biopsy is considered as a gold standard, there are limitations. Due to uneven distribution of NASH histologic features throughout the liver parenchyma, sampling errors could not be avoidable, resulting in substantial misdiagnosis and staging inaccuracies [32]. Also, liver biopsy is invasive and is not suitable for repeated evaluations within a short time period. Non-invasive, but accurate outcome measures, such as magnetic resonance elastography (MRE) and *Magnetic resonance imaging* derived proton density fat fraction (MRI-PDFF) are increasingly used in NAFLD/NASH trials to measure changes in hepatic elasticity or fat content assessing the entire liver.

The FDA guidance also provided some caveats for pediatric drug development. Pediatric NASH presents differently from adults and differs based on their pubertal development and sex [39]. As indicated in the FDA guidance, pediatric NASH also appears to have a different natural history when compared to adult NASH. Underlying mechanisms explaining the distinct NASH manifestation in children and adolescents remain to be investigated. However, since the histologic profile becomes similar to the one in adults as the pubertal development advances and some transitions are sex-specific, drastic changes in physiological sex hormone levels during the pubertal development might be involved in the histologic transition [39]. Such distinct histologic findings in paediatric NASH may indicates diverse disease mechanisms and repair process, requiring a special consideration in paediatric clinical trials. For reasons that are currently unknown, disease characteristics and progression in pediatric patients may be different.

4.4. Accelerated/Conditional Marketing Approvals

In practice, FDA has adopted policies to expedite drug development for serious medical conditions where few or no therapies exist. This is an evolving process which began in the 1980's with drugs intended to treat HIV, and continues through the present with multiple regulatory initiatives to facilitate availability of new drugs to patients – these programs either tackle the issue of accelerating clinical development and/or accelerating reviews by the regulatory agencies. For example, in 2014, FDA issued a guidance describing "Expedited Programs for Serious Conditions" which consolidated information on Fast Track, Breakthrough Therapy, Accelerated Approval, and Priority Review (FDA, 2014). Similarly, in 2014, the EMA launched *Adaptive Licensing* which, like the US *Accelerated Approval Pathway*, is also aimed at accelerating marketing approvals for products throughout the European Union (EMA, 2014). For development of products to treat NASH these accelerated pathways offer the potential to use surrogate endpoints to obtain *Accelerated* (US) and *conditional* (EU) marketing approval, with full marketing approval being granted with subsequent confirmatory studies using well established, and well-defined clinical outcomes.

With a registration pathway involving a two-stage process – namely, conditional marketing approval using surrogate endpoints, followed

by full marketing approval using well-defined clinical outcomes drug development programs to treat NASH seem well suited to continuous adaptive clinical studies. Conditional/accelerated marketing authorization are not new concepts. For example, in the 1980s accelerated marketing approvals were granted for a number of drugs to treat HIV based on surrogate endpoints, with post approval studies being required for full approval. Indeed, in 1997 Sheiner's disruptively innovative paper describing a *learn-confirm* strategy suggested we move away from thinking in terms of separate Phase 1/2/3 studies [37]. Although Sheiner's paper focused on early studies, the concept has been applied across the development continuum – namely, initial studies to learn about the drug and later studies to confirm a positive benefit-risk profile. There are, however, a number of stakeholders to be considered when designing seamless adaptive studies (patients, regulators, payers), and clearly one size does not fit all medical situations. However, given the challenges in developing therapies for NASH, it seems reasonable to assume that seamless adaptive designs provide a *good fit* in terms of development for marketing authorizations, and that these adaptive designs will continue to evolve as we learn more about this emerging epidemic.

4.5. Clinical/Statistical Perspectives

Based on the nature and the complexity of critical liver diseases, some considerations from clinical/statistical perspectives are necessary considered for not only providing an unbiased and reliable assessment of the test treatment under investigation and achieving study objectives with a desired power, but also to shorten and/or speed up the development process. These considerations are briefly described below.

4.6. Clinical Perspectives

In critical liver disease clinical trials, the following practical and challenging issues from clinical perspectives should be considered in order to shorten and/or speed up the drug development process under investigation: (i) is it possible to make decision early, (ii) whether biomarkers of surrogate endpoints can be used rather than regular clinical endpoints, (iii) is it possible to have options for sample size re-estimation, (iv) is it possible to drop inferior arms, (v) is it possible to add additional arms, and (vi) is it possible to conduct response-adaptive randomization after interim analyses.

As mentioned earlier, it may take a long time to measure study endpoint in liver disease clinical trials such as hepatitis C virus (HCV) infection clinical study. In HCV clinical study, a sustained virologic response (SVR), defined as an undetectable HCV RNA level (< 10 IU/mL) at 72 weeks (including 48 weeks of treatment and 24 weeks of follow-up), is considered acceptable endpoint for evaluation of patients with HCV infection. Since SVR will take a long time to assess the treatment effect, it is not feasible to make critical decision early based on the endpoint of SVR. Thus, there is desirable seeking for a surrogate endpoint or biomarker with a much shorter duration in order to make critical decision early. At the same time, it is also desirable to have options for conducting interim analyses. The purpose for the interim analysis is not only to (i) verify that the surrogate endpoint or biomarker is predictive of regular clinical endpoint, but also to (ii) perform sample size re-estimation to make sure that the study will achieve the desired power at the end of the study provided that the observed treatment effect at interim preserves. If the intended clinical study is for dose finding, it is desirable to have the options to drop or modify inferior arms or to add additional arms at interim for selecting promising doses moving forward to the next phase of clinical development. Moreover, to increase the probability of success, it is preferable that a response-adaptive randomization can be considered after the review of interim data.

In practice, to account for these practical and/or challenging issues and at the same time to shorten or speed up the development process of the test treatment under investigation, it is suggested that a wish list should be developed by the principal investigator so that the intended clinical trial can take the wish list into consideration for achieving the study objectives in a more efficient and timely fashion.

5. Statistical Perspectives

To provide an unbiased and reliable assessment of test treatment for treating critical liver disease, the following practical and challenging issues related to study objectives, study endpoints, and/or target patient populations when utilizing an adaptive trial design should be considered from statistical perspectives: (i) strategy for controlling bias/variation as the result of using adaptive trial design, (ii) implementation of power analysis for sample size calculation, (iii) control of overall type I error rate, (iv) statistical methods that fully utilize all data collected from different stages for a combined analysis, (v) criteria for critical decision-making at interim, (vi) implementation of blinded analysis at planned interim analyses, and (vii) integrity and validity of the intended clinical trial.

In clinical trials utilizing the adaptive trial design, bias and/or variation inevitably occur especially after the adaptations made to the on-going trial. [9] classified the sources of these bias and/or variation into the categories of: (i) expected and controllable, (ii) expected but not controllable, (iii) unexpected but controllable, and (iv) unexpected and not controllable (random error). In practice, the development of a strategy to control/eliminate these sources of bias and/or variation is of necessity for an unbiased and reliable assessment of the treatment effect of test treatment under investigation. The hypotheses to be tested should be able to reflect the study objectives of the intended study. Under the null hypothesis, an appropriate test statistic can then be derived for power calculation for sample size required at the pre-specified level of significance.

If a multiple-stage seamless adaptive trial design is used, appropriate statistical methods that utilize all data collected from all stages should be derived for a valid combined analysis which addresses the study objectives at different stages. In case there are planned interim analyses for critical decision-making, it is suggested that an independent

data monitoring committee (IDMC) should be established to perform blinded or unblinded analyses before recommendations to critical decision-making can be made. The establishment of IDMC plays an important role in clinical trials utilizing an adaptive trial design. It not only assures the integrity and validity of the intended clinical trial, but also assists the principal investigator in making critical decisions as the trial continues. Consequently, it increases the probability of success of the intended trial.

5.1. Complex Innovative Design

As indicated by PDUFA VI, a complex innovative design is referred to as a study design involving complex adaptations, Bayesian methods, or other features requiring simulations to determine operating characteristics. Thus, complex innovative designs include: (i) n-of-1 trial design [26, 14], (ii) adaptive trial design [9, 20], and (iii) master protocol design [40]. For development of drug products treating patients with liver diseases such as NASH, the following trials are necessarily conducted (i) early phase trials/proof-of-concept, (ii) phase 2 dose ranging, (iii) phase 3 trials, and (iv) phase 4 post-marketing study. To speed up and shorten the development process,

a two (multiple)-stage seamless adaptive trial design that combine several (independent) separate studies into a single trial is often considered. These seamless adaptive designs could include, but are not limited to: (i) proof-of-concept/dose ranging adaptive trial design, (ii) phase 3/4 adaptive trial design, and (iii) phase 2/3/4 adaptive design.

5.2. Two-Stage Seamless Adaptive Trial Design

A two-stage seamless adaptive trial design is a design that combines two (independent) separate studies into a single study that can address study objectives of both separate studies. In practice, two-stage adaptive designs can be classified into four categories depending upon whether their study objectives and study endpoints at different stages are the same [5, 6, 8, 11]. These categories include: (i) design with same study objectives and study endpoints at different stages (SS), (ii) designs with same study objectives but different study endpoints at different stages (SD), (iii) designs with different study objectives but same study endpoints at different stages (DS), and (iv) designs with different study objectives and different study endpoints at different stages (DD). To provide a better understanding of these designs, (Table 2) provides a list of summary of these study designs.

Chow (2020a) referred to SS design as a "0-D design", while SD and DS designs are considered as a "1-D design". The DD design is referred to as a "2-D design". The number of "D" is an indication of the number of differences either in study objective or in study endpoint. In this article, we will refer to these designs as "k-D design", where $k = 0, 1, \text{or } 2.$ The greater k is, the more complicated and problematic the seamless adaptive design is.

5.3. Analysis of Two-Stage Seamless Adaptive Designs

Statistical analyses for the "k-D design", where $k = 0, 1, \text{or } 2$ are brief described below.

5.3.1. The "0-D Design" (SS Design): SS Design is a two-stage seamless adaptive design with the same study objective and same study endpoint at different stages, which is similar to typical group sequential design with a planned interim analysis. Thus, standard statistical methods such as MIP (method of individual p-values), MSP (method of sum of p-values), and MPP (method of product of p-values) for group sequential design can be directly applied [4]. It should be noted that if additional adaptations such as change in primary study endpoint or hypotheses after the review of interim data, the standard methods have to be modified for the control of the overall type I error rate.

5.3.1. The "1-D Design" (SD Design or DS Design): A "1-D design" could be an SD design or a DS design, and the statistical analyses for an SD design and a DS design are different. To have a valid statistical analysis, some assumptions are necessary made. For example, for an SD design (i.e., study objectives at different stages are the same but the study endpoints are different at different stages), it is assumed that study endpoint (e.g., a biomarker, a surrogate endpoint, or a clinical endpoint with a short duration) at the first stage is predictive of the study endpoint (i.e., regular clinical endpoint) at the second stage [12]. On the other hand, for a DS design (i.e., study objectives at different stages are different but the study endpoints at different stages are the same), we have to consider testing two sets of hypotheses at different stages [8].

5.3.2. The "2-D Design" (DD Design): For the DD design (i.e., both study objectives and study endpoints at different stages are different), the following primary assumption and consideration are necessarily made for obtaining a valid statistical test using different endpoints for achieving study objectives at different stages: (i) study endpoint at the first stage is predictive of the study endpoint at the second stage, and (ii) consider testing two sets of hypotheses at different stages. A typical example of a 2-D design that combines a phase 2b study for treatment selection and a phase 3 study for efficacy confirmation, is a clinical trial for evaluation of safety, tolerability and efficacy of a test treatment for patients with hepatitis C virus (HCV) infection. More details regarding statistical analysis of a 2-D design (i.e., both study objectives and study endpoints are different at different stages) can be found in [8].

5.4. Remarks

In the previous sections, we only consider 4 types of two-stage seamless adaptive designs based on whether study objectives and study endpoints at different stages are the same or not. In NASH clinical trials, however, it is much more complicated. As it can be seen from (Table 1), in addition to study objectives and study endpoints, the target patient populations at different stages could be different. In this case, we can further classify two-stage seamless adaptive designs into eight categories depending upon whether the target patient population is the same or not. In other words, (Table 2) can be expanded to (Table 3).

Table 1: Objectives, Endpoints and Populations in NASH Clinical Trials

Phase	Primary Endpoint	target Population
Early phase trials/proof of concept	Endpoints should be based on mechanism of drug. Reduction in liver fat with a sustained improvement in transaminases; Improvement in biomarkers of liver inflammation opotosis and/or fibrosis. Consider using improvement in NAS (ballooning and inflammation) and/or fibrosis	Ideal to enroll patients with biopsy-proven NASH but acceptable to enroll patients at high risk for NASH (i.e., evidence of fatty liver two components of the metabolic syndrome, evidence of liver stiffness by imagining).
Dose ranging/phase 2	Resolution of NASH without worsening of fibrosis; alternatively, improvement in disease activity (NAS)/improvement in ballooning/inflammation without worsening of fibrosis	Biopsy proven NASH and NAS ≥4 inclide patients with NASH and liver fibrosis. Include a sufficient number of patients with NASH and fibrosis stage 2/3 to inform phase 3.
Trials to support a marketing application: phase 3	Resolution of steatohepatitis and no worsening of fibrosis Improvement in fibrosis with no worsening of steatohepaptis. A Co-primary endpoint of the bove or depending on the mode of action, either one or the other can be used.	patients with biopsy-confirmed NASH with moderate/advanced fibrosis (F2/F3)
Trials to support a marketing application: phase 4(postmarketing part)	Clinical outcome trial underway by the time of submission; Composite endpoint histopathologic progression to circhosis; MELD score change > 2 points or MELD increase to > 15 in population enrolled with MELD ≤ 13; death; tranplant; Circhosis decompensation events;	Patients with biopsy-confirmed NASh with moderate/advanced fibrosis (F2/F3).

Table 2: Types of Two-Stage Seamless Adaptive Designs (Depending upon Objective and Endpoint)

	Study Endpoint	
Study Objectives	Same (S)	Different (D)
Same (S)	SS	SD
Different (D)	DS	DD

Table 3: Types of Two-Stage Seamless Adaptive Designs (Depending upon Objective, Endpoint, and Population).

	Target Patient Population			
	Same (S)		Different (D)	
	Study Endpoint		Study Endpoint	
Study Objective	Same (S)	Different (D)	Same (S)	Different (D)
Same (S)	SSS	SDS	SSD	SDD
Different (D)	DSS	DDS	DSD	DDD

(Table 3) indicates that there are one "0-D design", three "1-D designs", three "2-D designs", and one "3-D design" if we account for possible shift in patient population. If there is no shift in patient population from stage to stage, Table 3 reduces to (Table 2). In this case, statistical methods as described in the previous section can be applied. In case there is population shift from stage to stage, in addition to the primary assumption and consideration that are necessarily made for derivation of valid statistical test under specific "1-D design" or "2-D design", an evaluation of potential shift in patient population from stage to stage need to be considered. For this purpose, we may consider the method proposed by [29] for assessment of possible population shift from stage to stage by evaluating the sensitivity index under different models that (i) shift in location parameter is random, (ii) shift in scale parameter is random, and (iii) shifts in both location and scale parameters are random.

6. Case Studies

6.1 Case Study #1 – The HCV Dose Selection Study

A pharmaceutical company is interested in conducting a dose selection clinical trial for evaluation of a test treatment as compared to a standard care (SOC) in treating subjects with hepatitis C virus (HCV) genotype 1 infection. After consulted with FDA reviewers, the sponsor planned to conduct a phase 2 study for dose finding and a phase 3

study for efficacy confirmatory in order to fulfill with FDA's requirement for regulatory submission. The sponsor was interested in shortening the development process. Thus, it is suggested that a single trial with two-stage seamless adaptive trial design should be conducted for achieving the study objectives of phase 2 dose finding, and phase 3 efficacy confirmation of the test treatment under investigation. As a result, the proposed two-stage seamless adaptive trial design consists of two stages: dose selection and efficacy confirmation. For a more detailed description of the study design, see [8].

For efficacy confirmation, the primary efficacy endpoint, which is accepted to the regulatory agency, is the incidence of sustained virologic response (SVR), defined as an undetectable HCV RNA level (< 10 IU/mL) at 24 weeks after treatment is complete (Study Week 72). In the interest of shortening the development process, the same endpoint with a much shorter duration is proposed for dose finding. In other words, the study endpoint of early virologic response (EVR) at week 12 is considered. In both Stages 1 and 2, consider the incidence of the following primary efficacy variables of (i) rapid virologic response (RVR), that is, undetectable HCV RNA level at Study Week 4; (ii) early virologic response (EVR), that is, ≥ 2-log10 reduction in HCV RNA level at Study Week 12 compared with the baseline level; (iii) end-of-treatment response (EOT), that is, undetectable HCV RNA level at Study Week 48; and (iv) SVR, that is, undetectable HCV RNA level at Study Week 72 (24weeks follow up after treatment is complete). Thus, the study objectives at different stages are similar but different, while the study endpoints are different. Thus, the proposed two-stage seamless adaptive trial design is a typical 2-D design.

Suppose there are a total of four treatments including three active treatments (doses) and one placebo and two planned interim analyses. The proposed two-stage seamless adaptive trial design is briefly described below.

Stage 1 is a four-arm randomized evaluation of three dose levels of continuous subcutaneous (SC) delivery of the test treatment compared with pegylated interferon α (standard of care) given as once weekly SC injections. All subjects will receive oral weight-based ribavirin. After all Stage 1 subjects have completed Study Week 12, an interim analysis will be performed. The interim analysis will provide information to enable selection of an active dose of the test treatment based on safety/tolerability, outcomes, and early indications of efficacy to proceed to testing for non-inferiority compared with standard of care in Stage 2. Depending upon individual response data for safety and efficacy, Stage 1 subjects will continue with their randomization assignments for the full planned 48 weeks of therapy, with a final follow-up evaluation at Study Week 72. Stage 2 will be a non-inferiority comparison of a selected dose and the same pegylated interferon α active-control regimen used in Stage 1, both again given with oral ribavirin, for up to 48 weeks of therapy, with a final follow-up evaluation at Study Week 72. A second interim analysis of all available safety/tolerability, outcomes, and efficacy data from Stage 1 and Stage 2 will be performed when all Stage 2 subjects have completed Study Week 12. Depending upon individual response data for safety and efficacy, Stage 2 subjects will receive the full planned 48weeks of treatment, with final follow-up at Study Week 72. A diagram of two-stage seamless adaptive trial design is illustrated in (Figure 1) (see also [8].

For sample size estimation, a total of 388 subjects are required (i.e., 120 subjects or 30 subjects per arm for Stage 1 and 268 subjects or 134 subjects per arm for Stage 2) to account for a probable dropout rate up to 15% as well as 2 planned interim analyses using the O'Brien-Fleming method. Stage 1 enrolled a total of 120 subjects at 1:1:1:1 treatment allocation ratio among the four treatment arms. The purpose was to gather sufficient data for selection of a promising active dose proceeding to Stage 2. Stage 2 enrolled an additional cohort of 268 subjects at 1:1 treatment allocation ratio between two treatment arms. The purpose was to provide a sufficient number of subjects to establish non-inferiority of continuous interferon delivery to standard-of-care interferon therapy, for a combined analysis based on a total of 306 subjects from Stage 1 and Stage 2. For this purpose, it was expected that 164 subjects from the selected dose arms (30 from Stage 1 plus 134 additional subjects enrolled in Stage 2) and 164 subjects from the pegylated interferon α active-control arms (30 from Stage1 plus an additional 134 subjects enrolled in Stage 2) are required to meet the study objective of achieving an 80% power for establishing non-inferiority (with a non-inferiority margin of 15%) at the 5% level of significance (i.e., an overall type I error rate of 5%). Detailed information regarding statistical methods for data analysis can be found in [8].

Note that above two-stage seamless trial design is a typical complex innovative design which consists of the concepts of (i) group sequential design, (ii) drop-the-losers design, and (iii) seamless adaptive phase 2/3 design utilizing precision analysis (i.e., confidence interval approach) for decision-making on dose selection at the first stage. If we apply adaptive randomization for the second stage, the study design would be even more complicated. From regulatory point of view, it is important to ensure that (i) no operational biases are introduced during the conduct of the trial and (ii) the overall type I error rate is well controlled at the 5% level of significance.

6.2. Case Study #2 – Clinical Development NASH Program

For development of drug products for treating patients with NASH, after having consulted with regulatory agency, it is suggested the following clinical trials utilizing seamless adaptive designs may be useful to shorten and speed up the process of NASH drug product development: (i) proof-of-concept/dose ranging adaptive trial design, (ii) phase 3/4 adaptive trial design, and (iii) phase 2/3/4 adaptive design. For illustration purpose, consider a single seamless phase 2/3/4 adaptive trial design allows adaptations, continuous exposure, and long-term follow-up (see Figure 2).

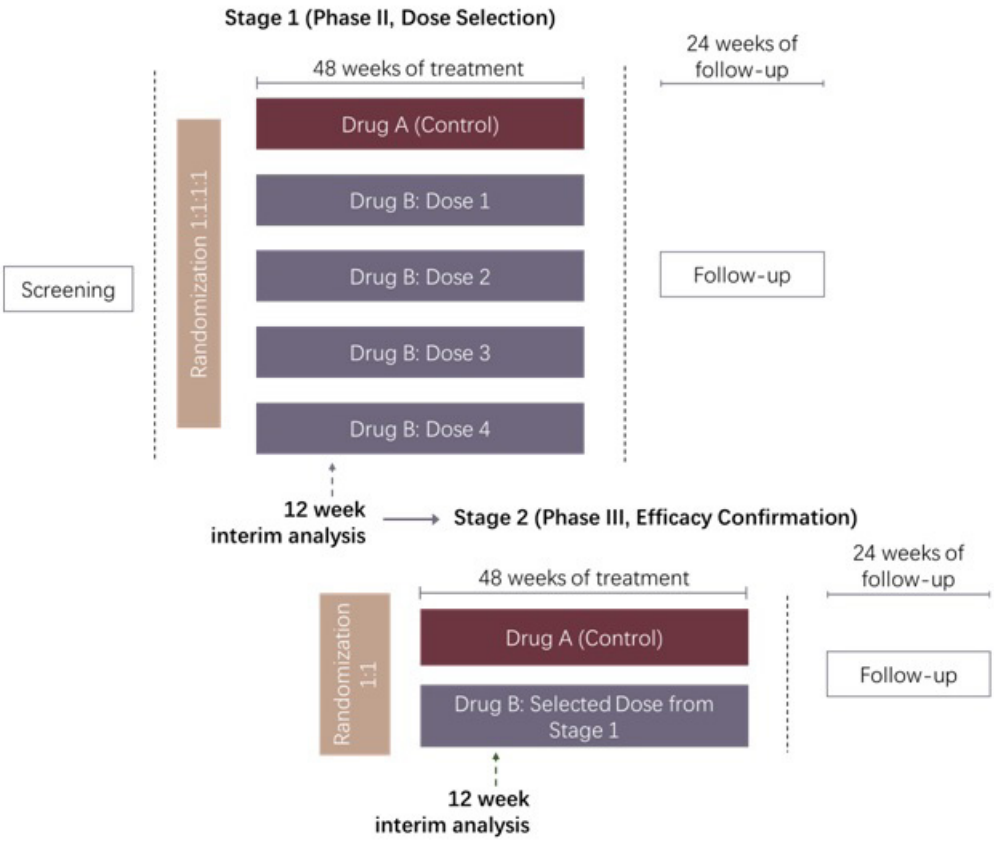

Figure 1: A typical HCV clinical trial utilizing a phase 2/3 seamless adaptive trial design

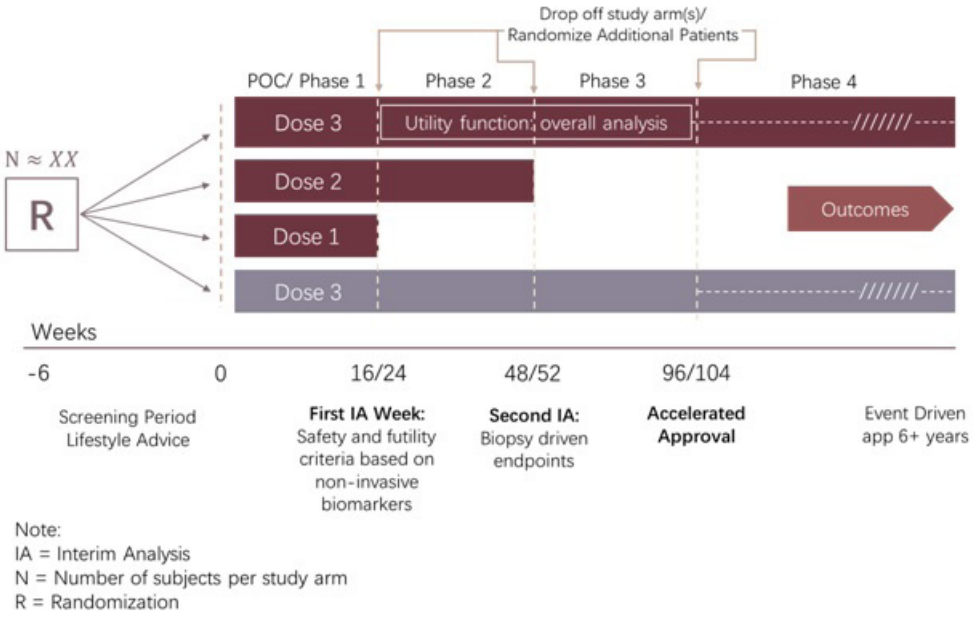

Figure 2: A typical NASH clinical study utilizing a phase 2/3/4 seamless adaptive trial design.

As discussed earlier, a NASH clinical study utilizing a phase 2/3/4 seamless adaptive trial design is a typical "3-D" design which involves different study objectives, different study endpoints, and different target patient populations (due to progression of the disease) at different stages (see Table 1). The phase 2/3/4 seamless adaptive trial design combines a phase 2, phase 3, and phase 4 into a single study. The study objectives of phase 2, phase 3, and phase 4 are for dose finding, efficacy confirmation, and in support of marketing application, respectively, which are different stages of the phase 2/3/4 seamless adaptive clinical trial. At the first two stages (i.e., stage 1 for phase 2 trial and stage 2 for phase 3 study), study endpoints such as (i) reduction of at least 2 points in NAS, (ii) resolution of NASH by histology without worsening of fibrosis, and/or (iii) improvement in fibrosis without worsening of NASH are considered. At the third stage, a post-marketing phase 4 with demonstration of improvement in clinical outcomes will lead to final marketing authorization. The proposed three-stage phase 2/3/4 seamless adaptive trial design is illustrated in (Figure 2).

A single phase 2/3/4 seamless adaptive trial design allows adaptations, continuous exposure, and long-term follow-up. Study endpoints at interim analysis are reduction of at least 2 points in NAS, resolution of NASH by histology without worsening of fibrosis, and/or improvement in fibrosis without worsening of NASH. One or two doses (the most promising doses) may continue to the next stage. A post-marketing phase 4 with demonstration of improvement in clinical outcomes will lead to final marketing authorization.

Although the above seamless phase 2/3/4 appears to be reasonable, regulatory agency such as FDA emphasizes that the designs must be supported by a sound rationale and scientific justifiable for integrity, quality and validity. Thus, it is suggested that a study protocol for a NASH clinical trial utilizing the proposed phase 2/3/4 seamless adaptive trial design should be developed for addressing the following typical issues:

- Provide detailed information regarding how the overall type I error rate is controlled or preserved;
- Provide a detailed strategy or plan for preventing possible operational biases that may be incurred before and after the adaptations are applied;
- Provide justification regarding the validity of statistical methods used for a combined analysis;
- Provide justification for the chosen alpha spending function (e.g., O'Brien-Fleming) for stopping boundaries;
- Provide justification regarding criteria used for critical decision-making at interims;
- Establish an independent data safety monitoring committee (IDMC) and provide IDMC charter;
- Provide justification for power analysis for sample size calculation and sample size allocation especially where the study objectives, endpoints, and populations are different at different stages;
- Provide justification if sample size re-estimation is performed in a blinded or unblinded fashion in the seamless adaptive trial design.

Note that since only one single trial would lead to regulatory approval, a very small overall alpha (e.g., <0.001) is recommended to ensure proper control of the overall type error rate.

7. Concluding Remarks

For development of treatment for chronic liver diseases such as HBV/HCV and NAFLD/NASH, it will generally take a long time (e.g., more than 10-20 years) to conduct. In practice, it is not feasible to conduct such clinical studies. Besides, liver disease clinical trials may involve different study objectives, different study endpoints, and/or different target patient populations depending upon the stages of the progression of the disease. To account for these complications, regulatory agencies such as FDA suggests the use of complex innovative design to (i) provide an unbiased and reliable assessment of the treatment effect of the test drug under investigation at different stages of disease progression, (ii) shorten the development process, and (iii) speed up the overall path to regulatory review and approval.

In this article, several complex innovative designs such as multiple-stage seamless adaptive designs are introduced for liver disease clinical trials such as HBV/HCV and NAFLD/NASH clinical trials. Under a multiple-stage seamless adaptive trial design, when there are differences in study objective, study endpoint, and/or target patient population, it is suggested that the following be considered: (i) testing two sets of hypotheses (when study objectives are not the same at different stages), (ii) the assumption that the study endpoint at the first stage is predictive of the study endpoint at the second stage (when the study endpoints are different at different stages), and/or (iii) the assessment of possible population shift need to be taken into consideration (when there are population shift from stage to stage due to disease progression). Under these assumptions and/or considerations, a valid statistical test for addressing the scientific/medical questions of interest can be obtained.

Regarding the development of drug products for chronic liver diseases such as NASH, one of the greatest challenges currently faced, prior to enrolling patients into NASH clinical trials, is deciphering which patients with NAFLD have NASH, particularly those with advanced fibrosis. Once these at-risk patients have been identified, the endpoints that appear to be the most readily attainable and reliable include monitoring for fibrosis regression, development of cirrhosis, and surrogate measures of liver-related outcomes. Longer-term follow-up to assess for all-cause mortality (mainly cardiovascular death) and liver-related mortality is also important but will take longer to evaluate. Developing novel, non-invasive technology to assess these endpoints is imperative to achieve global success in finding effective

therapies for NASH [22]. The development and acceptance of meaningful, readily obtainable, and well-defined clinical trial endpoints in clinical liver diseases such as HBV/HCV and NAFLD/NASH are imperative to develop new and effective therapies to treat this growing epidemic. To address this issue, [21] proposed a therapeutic index function for analysis of two-stage seamless adaptive designs with distinct study endpoints at each stage under the assumption that even though the two endpoints are not the same, there is well-established relationship between them. A therapeutic index function is defined for each one of the endpoints (often consists of a number of criteria). It takes different endpoints with pre-specified criteria into consideration and it is based on a vector of therapeutic index function rather than individual endpoints. The vector of therapeutic index model allows the investigator to accurately and reliably assess the treatment effect in a more efficient way [7].

References

1. Angulo P, Kleiner DE, Dam-Larsen S, Adams LA, Bjornsson ES, Charatcharoenwitthaya P et al. Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with non-alcoholic fatty liver disease. Gastroenterology. 2015; 149: 389-397.e10.
2. Argo CK, Northup PG, Al-Osaimi AM, Caldwell SH. Systematic review of risk factors for fibrosis progression in non-alcoholic steatohepatitis. J Hepatol. 2009; 51: 371-379.
3. Brunt EM, Kleiner DE, Wilson LA, Belt P, Neuschwander-Tetri BA. NASH Clinical Research Network (CRN). Nonalcoholic fatty liver disease (NAFLD) activity score and the histopathologic diagnosis in NAFLD: distinct clinicopathologic meanings. Hepatology. 2011; 53: 810-820.
4. Chang M. Adaptive design method based on sum of p-values. Statistics in Medicine. 2007; 26: 2772-84.
5. Chow SC. Complex innovative design for NASH clinical trials. Academic Journal of Gastroenterology & Hepatology. 2020a: DOI: 10.33552/AJGH.2020.02.000538
6. Chow SC. Innovative Methods for Rare Disease Drug Development. Chapman and Hall/CRC Press, Taylor & Francis, New York. 2020b.
7. Chow SC, Huang Z. Innovative thinking on endpoint selection in clinical trials. Journal of Biopharmaceutical Statistics. 2019; 29: 941-951.
8. Chow SC, Lin M. Analysis of two-stage adaptive seamless trial design. Pharmaceutica Analytica Acta. 2015; 6: 341.
9. Chow SC, Chang, M. Adaptive Design Methods in Clinical Trials. 2nd edition, Chapman and Hall/CRC Press, Taylor & Francis, New York, New York. 2011.
10. Chow SC. and Corey R. Benefits, Challenges and obstacles of adaptive designs in clinical trials. The Orphanet Journal of Rare Diseases. 2011; 6: 79 doi:10.1186/1750-1172-6-79.
11. Chow SC, Tu, YH. On two-stage seamless adaptive design in clinical trials. Journal of Formosan Medical Association. 2008; 107: No.12, S51-S59.
12. Chow SC, Lu Q, Tse SK. Statistical analysis for two-stage adaptive design with different study endpoints. Journal of Biopharmaceutical Statistics. 2007; 17: 1163-1176.
13. Cvitanović Tomaš T, Urlep Ž, Moškon M, Mraz M, Rozman D et al. LiverSex Computational Model: Sexual Aspects in Hepatic Metabolism and Abnormalities. Front Physiol. 2018; 9: 360. doi: 10.3389/fphys.2018.00360. PMID: 29706895; PMCID: PMC5907313.
14. Davidson KW, Cheung YK, McGinn T, Wang YC. Expanding the role of n-of-1 trials in the precision medicine era: actionpriorities and practical consideration. Commentary. National Academy of Medicine, Washington DC. 2018. http://doi.org/10.31478/20181 2d.
15. Ekstedt M, Franzen LE. Mathiesen UL, Thorelius L, Holmqvist M, Bodemar G et al. Long-term follow-up of patients with NAFLD and elevated liver enzymes. Hepatology. 2006; 44: 865-873.
16. EMA. Pilot project on adaptive licensing. European Medicines Agency, London, UK. 2014. https://www.ema.europa.eu/en/documents/other/pilot-project-adaptive-licensing_en.pdf
17. FDA. Guidance for Industry – Expedited Programs for Serious Conditions – Drugs and Biologics. The United States Food and Drug Administration, Silver Spring, Maryland. 2014.
18. FDA. Guidance for Industry – Chronic Hepatitis B Virus Infection: Developing Drugs for Treatment. The United States Food and Drug Administration, Silver Spring, Maryland. 2018a.
19. FDA. Guidance for Industry – Noncirrhotic Nonalcoholic Steatohepatitis with Liver Fibrosis: Developing Drugs for Treatment. The United States Food and Drug Administration, Silver Spring, Maryland. 2018b.
20. FDA. Guidance for Industry – Adaptive Designs for Clinical Trials of Drugs and Biologics. The United States Food and Drug Administration, Silver Spring, Maryland, November. 2019.
21. Filozof C, Chow SC, Dimick-Santos L, Chen YF, Williams RN, Goldstein BJ et al. Clinical endpoints and adaptive clinical trials in precirrhotic nonalcohotic steatohepatitis: facilitating development approaches for an emerging epidemic. Hepatology Communications. 2017; 1: 577-585.
22. Hannach WN, Torres DM, Harrison SA. Nonalcoholic steatohepatitis and endpoints in clinical trials. Gastroenterology & Hepatology. 2016; 12: 756-763.
23. Kim IH, Kisseleva T, Brenner DA. Aging and liver disease. Curr Opin Gastroenterol. 2015; 31:184-91. doi: 10.1097/MOG.0000000000000176. PMID: 25850346; PMCID: PMC4736713.
24. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW et al. Nonalcoholic Steatohepatitis Clinical Research Network. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology. 2005; 41: 1313-21.
25. Lauer GM, Walker BD. Hepatitis C virus infection. New England Journal of Medicine. 2001; 345: 41-52.
26. Lillie EO, Patay B, Diamant J, Issell B, Topol E, Schork NJ. Then-of-1 clinical trial: the ultimate strategy for individualizing medicine? Personal Med. 2011; 8: 161–73. https ://doi.org/10.2217/pme.11.7.

27. Lonardo A, Nascimbeni F, Ballestri S, Fairweather D, Win S, Then TA, Abdelmalek MF, Suzuki A et al. Sex Differences in Nonalcoholic Fatty Liver Disease: State of the Art and Identification of Research Gaps. Hepatology. 2019; 70: 1457-69. doi: 10.1002/hep.30626. Epub 2019 Sep 23. PMID: 30924946; PMCID: PMC6766425.

28. Lonardo A, Suzuki A. Sexual Dimorphism of NAFLD in Adults. Focus on Clinical Aspects and Implications for Practice and Translational Research. J Clin Med. 2020; 9: 1278. doi: 10.3390/jcm9051278. PMID: 32354182; PMCID: PMC7288212.

29. Lu Y, Kong YY, Chow SC. Analysis of sensitivity index for assessing generalizability in clinical research. Jacobs Journal of Biostatistics. 2017; 2: 009.

30. Musso G, Gambino R, Cassader M, Pagano G. Meta-analysis: natural history of non-alcoholic fatty liver disease (NAFLD) and diagnostic accuracy of non-invasive tests for liver disease severity. Annals of Medicine. 2011; 43: 617-649.

31. Pagadala MR, McCullough AJ. The relevance of liver histology to predicting clinically meaningful outcomes in nonalcoholic steatohepatitis. Clin Liv Dis. 2012; 16: 487-504.

32. Ratziu V, Charlotte F, Heurtier A, Gombert S, Giral P, Bruckert E, Grimaldi A et al. LIDO Study Group. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology. 2005; 128: 1898-906. doi: 10.1053/j.gastro.2005.03.084. PMID: 15940625.

33. Rehermann B, Nascimbeni M. Immunology of hepatitis B virus and hepatitis C virus infection. Immunology. 2005; 5: 215-229.

34. Sanyal AJ. NASH: A global health problem. Hepatol. Res. 2011; 41: 670-674.

35. Sanyal A, Brunt E, Kleiner D, Kowdley KV, Chalasani N, Lavine JE et al. Endpoints and clinical trial design for nonalcoholic steatohepatitis. Hepatology. 2011; 54: 344-353.

36. Sanyal AJ, Friedman SL, McCullough AJ, Dimick-Santos L, American Association for the Study of Liver Diseases, & United States Food and Drug Administration. Challenges and opportunities in drug and biomarker development for nonalcoholic steatohepatitis: findings and recommendations from an American Association for the Study of Liver Diseases-U.S. Food and Drug Administration Joint Workshop. Hepatology. 2015; 61: 1392–1405.

37. Sheiner L. Learning versus confirming in clinical drug development. Clin Pharmacol Ther. 1997; 61: 275–291.

38. Singh S, Allen AM, Wang Z, Prokop LJ, Murad MH, Loomba R et al. Fibrosis progression in nonalcoholic fatty liver vs nonalcoholic steatohepatitis: a systematic review and meta-analysis of paired biopsy studies. Clin Gastroenterol Hepatol. 2015; 13: 643-54. e1-9.

39. Suzuki A, Abdelmalek MF, Schwimmer JB, Lavine JE, Scheimann AO, Unalp-Arida A et al. Nonalcoholic Steatohepatitis Clinical Research Network. Association between puberty and features of nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol. 2012; 10: 786-94. doi: 10.1016/j.cgh.2012.01.020. Epub 2012 Feb 14. PMID: 22343513; PMCID: PMC3382041.

40. Woodcock, LaVange, L.M. Master protocols to study multiple therapies, multiple diseases, or both. The New England Journal of Medicine. 2017; 377: 62-70.

41. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M et al. Global epidemiology of non-alcoholic fatty liver disease-meta-analytic assessment of prevalence, incidence and outcomes. Hepatology. 2016; 64: 73-84.

42. Younossi ZM, Stepanova M, Afendy M, Fang Y, Younossi Y, Mir H et al. Changes in the prevalence of the most common causes of chronic liver diseases in the United States from 1988 to 2008. Clin. Gastroenterol. Hepatology. 2011; 9: 524-530. e1.