A critical review of endpoints for non-cirrhotic NASH therapeutic trials

Vlad Ratziu*

Summary

Non-alcoholic steatohepatitis is a disease without a single, specific, diagnostic marker, hence multiple indicators are required to measure therapeutic efficacy. Moreover, drug candidates for non-alcoholic steatohepatitis target many distinct mechanisms that are believed to promote hepatic injury. Therefore, a wide range of endpoints must be reached, sequentially, as required by the drug development process. Some of these endpoints validate the mechanism of action, others are used to anticipate histological efficacy. Histological endpoints are still considered the best predictors of clinical outcome, but they can only be reliably tested in larger, late phase trials. Herein, we will review the rationale and clinical data supporting the use of specific endpoints at different stages of therapeutic trials. We will also discuss the validity and limitations of current phase IIb histological endpoints, particularly a one stage reduction in fibrosis, for their ability to predict progression to cirrhosis, which is the ultimate outcome measure in therapeutic trials.

© 2017 Published by Elsevier B.V. on behalf of European Association for the Study of the Liver.

Introduction

Therapeutic trials need to provide answers to precise questions and these questions are different at the various stages of drug development. Early phase trials (phase I and Ia) are designed to provide information about pharmacokinetic parameters, early and short-term human safety data and pharmacodynamics. These are common to most therapeutic fields or classes of drugs with little if any specificity for particular diseases, such as non-alcoholic fatty liver disease (NAFLD) and non-alcoholic steatohepatitis (NASH). In addition, phase Ila trials also demonstrate "proof of principle", meaning measurable “on-target effects” and their biological consequences. A particularly important outcome of phase Ila trials is the selection of one or two doses that provide maximum efficacy with acceptable safety and tolerability. This provides precious insight into the dose dependency of the therapeutic effect and the safety margin of a particular compound. These doses will be carried over to later stage trials. Typically phase IIb trials will explore whether the biological effects observed in earlier trials translate into hepatic histological improvement. Finally, large scale phase III and subsequent outcome trials will provide a comprehensive and statistically robust demonstration of the benefit in terms of histological improvement and long-term clinical outcomes. Since each trial will have to deliver information critical for designing the next step, it is necessary to carefully consider how the endpoints should be chosen and if they can be achieved within a given time frame.

Early phase trials

While endpoints in phase I studies are standard and will not be further discussed herein, those for proof of principle, phase Ila trials are usually defined based on the mechanism of action of each drug. The drugs currently in development for NASH can be divided into three broad categories: metabolic, anti-inflammatory and antifibrotic. The crucial question is, if and how do the metabolic effects of a drug candidate translate into improvements in hepatic histology? The same question must be asked for the potential anti-inflammatory and antifibrotic properties since, with few exceptions, most early phase trials do not use direct histological assessments.

Metabolic effects

Since NASH is a disease intimately associated with insulin resistance, adipose tissue dysfunction and the various phenotypic manifestations of the metabolic syndrome, (mainly overweight, visceral adiposity, dyslipidaemia, arterial hypertension and glucose abnormalities) a wide variety of clinical (weight changes) and biological variables can be measured to determine how the drug affects these underlying metabolic abnormalities. Weight reduction is the most straightforward indicator of a possible benefit for NASH:1 with diet and lifestyle interventions there is evidence that modest (<5% from baseline) weight loss can reduce steatosis, higher levels (5–8%) can improve hepatic inflammation ballooning and clear NASH, and more marked weight loss (>10%) can even reduce fibrosis.2 Although data are currently lacking, it is probable that pharmacologically induced weight loss of the same magnitude will result in the same histological effects. However, most drugs are either weight neutral or, as in the case of glitazones, increase weight while still improving the liver. Therefore, again, the mechanism of action defines the outcomes and the same information only has value within a particular context.

Keywords: Randomised controlled trials; NASH; Fibrosis; Steatohepatitis; Steatosis; Magnetic resonance imaging; Trial outcomes.

Hospital Pitié-Salpêtrière, Institute for Cardiometabolism and Nutrition and Université Pierre et Marie Curie, Sorbonne Universités, Paris, France

Key point

Endpoints for phase Ila trials are usually defined based on a drug’s mechanism of action. Three broad categories of drugs currently in development for NASH are metabolic, anti-inflammatory and antifibrotic.

* Address: Hospital Pitié-Salpêtrière, 47-83 bd de l’Hôpital, Paris 75013, France. E-mail address: Vlad.ratziu@inserm.fr.
Accurate quantification of insulin action in individuals with diabetes and of tissue/pathway specific insulin resistance necessitates the use of cumbersome techniques, such as the euglycaemic insulin clamp combined with infusion of labelled glucose or free fatty acids. These complex methods are rarely, if ever, conducted in liver clinics. Instead, improvement in insulin resistance can be accurately demonstrated in non-diabetic individuals by surrogate measures, such as homeostatic model assessment of insulin resistance (fasting glucose multiplied by insulin) [5-7] concentrations of serum adiponectin, or the adipose tissue insulin resistance index (the product of free fatty acids and fasting insulin). Like with all surrogates, the question is, to what extent is an improvement in insulin resistance predictive of an improvement in hepatic histology?

Trials with glitazones have provided the best evidence so far that improvement in insulin sensitivity can be associated with histological improvement.[11-13] Both pioglitazone and rosiglitazone have a strong anti-steatogenic effect,[12,14-16] which is possibly mediated by rises in adiponectin levels and a reduction of adipose tissue-derived lipolysis,[10,17] Since the action of insulin on sterol regulatory element-binding protein 1c and consequently on several of the key enzymes of lipogenesis remains intact in insulin resistance states,[18] there is little evidence that an insulin sensitiser drug would change the rate of lipogenesis. Clinical data suggest that an improvement in insulin sensitivity coexists with an improvement in hepatic necroinflammation; pioglitazone has been shown to resolve NASH more often than placebo.[12] This could be due to a reduction in lipotoxic precursors because of a better control of lipolysis, or to increases in adiponectin which may have anti-inflammatory properties.[19] However, much more work is necessary to understand both the clinical reality of this link and its biological determinants. Whether an improvement in insulin resistance results in an improvement in hepatic fibrosis is unknown. Clinical results alone are inconclusive, no matter what meta-analyses are tempted to conclude.[20] No studies designed to assess fibrosis improvement are available, let alone large scale or longer-term trials. There is of course the possibility of an indirect effect, since improving the conditions that created NASH may subsequently shut-off the fibrogenic process. This could also be facilitated by the silencing of some necroinflammatory hepatic damage, which could then inhibit the triggers of fibrosis. There are also “off-target” effects since, for instance, stellate cells express several nuclear transcription factor receptors, such as peroxisome proliferator-activated receptor (PPAR)γ. PPARγ transcriptional activity is reduced during the activation of stellate cells into a myofibroblastic phenotype,[21-23] and activation of PPARγ reduces[21,22] to some extent hepatic fibrosis. However, more work is needed to understand if there is a more direct effect of insulin sensitisation on the myofibroblastic activation of stellate cells that could be mediated through many mediators including adiponectin.[19,25] As far as the effect of a drug on the phenotypic manifestations of the metabolic syndrome, there are multiple standard fasting or dynamic parameters (oral glucose tolerance tests or lipid load tests)[26] of glucose homeostasis and lipid alterations that can be measured. Favourable metabolic effects of a drug are desirable in patients with NASH but not mandatory. To what extent the partial correction of these metabolic abnormalities is predictive of improvements in hepatic histology is unknown. Therefore, these outcomes are primarily useful to validate the pharmacodynamics and biological actions of a drug and less for the prediction of histological improvement.

Metabolic improvement (either through weight reduction or enhanced insulin sensitivity) results in a reduction in steatosis. Imaging modalities for a precise quantification of steatosis are now available using magnetic resonance spectroscopy[27] and MRI-based proton density fat fraction (PDFF).[28,29] The latter has the advantage of being easier to implement, even in multicentric studies on different MRI machines. The physiological amount of liver fat in a healthy population, as measured by these imaging methods, is around 5.5%.[30] Fat quantification by MRI-PDFF is strongly correlated with the histological semi-quantitative assessment of steatosis and its changes over time.[31] Paired histology-MRI-PDFF data from a subset of the FLINT trial has shown that a 5-6% absolute change in PDFF correlates with steatosis improvement or aggravation as measured by liver biopsy.[32] Another study has shown the same correlation with histology for a 30% relative change.[33] Several phase IIa proof of principle studies now use steatosis quantification by MRI-PDFF to detect short-term changes in liver fat after therapy.[34-36]

Anti-inflammatory and antifibrotic effects

There are no good serum or imaging markers of steatohepatitis, improvement in liver cell injury or cell death, or the hepatic inflammatory cascade. CK-18 fragments are a rather unreliable marker of steatohepatitis[37] and changes in CK-18 on therapy are at best difficult to interpret, even when using anti-apoptotic molecules.[38] A reduction in soluble markers of systemic inflammation such as interleukin (IL)-1β, IL-6, high sensitivity C-reactive protein or fibrinogen can result from enhanced insulin sensitivity or reduced adipose tissue insulin resistance, but evidence of a direct correlation with improvement in hepatic inflammation is lacking.[39,40] It is therefore questionable whether these circulating inflammatory markers are a reliable indicator of changes in hepatic inflammation. Thus, their value as an outcome measure is simply
to validate the metabolic or systemic effects and not necessarily the effects on hepatic injury.

Considering these limitations, we are currently left with the measurement of serum aminotransferases as the only non-invasive marker of the hepatic anti-inflammatory effects of a drug. Most, if not all studies that have shown clear improvements in hepatic histology, have also documented a robust decrease in alanine aminotransferase (ALT) levels. A robust decrease is in the range of 30–40%, which is maintained throughout the entire treatment period. ALT values usually do not change in the placebo arm, although in some studies/populations there could be an initial decline. However, a sustained decline of ALT values in the placebo arm is unusual and should raise suspicion about major diet or lifestyle changes during the trial. Also, some drugs have failed to reduce ALT levels, somehow giving credibility to this biochemical signal. Whether a positive signal (ALT reduction) always predicts histological improvement is unknown. Despite lingering methodological controversies, ursodeoxycholic acid is an interesting example of a drug which has led to a strong reduction in ALT values in some NASH studies (and also in chronic hepatitis C), but no histological benefit in other studies. Conversely, most researchers consider a lack of ALT change as indicative of drug inefficacy, and hence a no-go signal for further testing of anti-inflammatory drug candidates. This may be true for improvement in steatohepatitis, but it is unclear whether it applies to antifibrotic drugs as well. There is at least one recent example where a drug claiming an antifibrotic effect did not change ALT values.

Antifibrotic effects are even harder to assess without histology, particularly in short-term phase IIa trials. Both fibrosis build-up and fibrosis reversal are slowly evolving processes and it may take years after removal of the primary disease before a meaningful reduction in fibrosis becomes detectable. One may ask what the rationale for looking for fibrosis changes in 12-week phase IIa trials is, other than trying to understand if there is any chance of a direct antifibrotic effect of a drug. Direct fibrogenic markers could be the most useful for this purpose, but pro-C3 or matrix remodelling rates have not been sufficiently validated. What is clear is that standard serum fibrosis markers (FibroTest, NFS [NAFLD fibrosis score], enhanced liver fibrosis [ELF]) or elastometry can rapidly decline in the absence of a reduction in fibrosis. For instance, inflammation and ALT levels influence liver stiffness beyond the amount of fibrotic scarring and may confound early changes in elastometry values. The same could hold true for serum fibrosis markers, as studies in patients treated for viral hepatitis have shown early changes in serum levels, before changes in fibrosis were documented. Thus, the predictive value of serum markers and elastometry for fibrotic changes lack both specificity and most probably sensitivity in early phase trials. A lot of hope is placed on 2D magnetic resonance elastography technology, which could be a valuable imaging marker of fibrosis in cross sectional studies. It has been suggested that a 15% relative reduction in stiffness, measured by magnetic resonance elastography, represents a one stage reduction in fibrosis measured by histology. Unfortunately data from clinical trials testing multiple doses of the same drug are not yet entirely demonstrative, although more data will be available in the near future. Recently it was suggested that a T1 mapping technique for fibrosis and inflammation using multiparametric magnetic resonance imaging could be a promising method to grade the severity of steatohepatitis and to predict clinical events. However, these small studies await independent confirmation from larger trials.

**Late phase trials**

**Overall objectives of phase IIb trials**

Phase IIb trials are typically performed on hundreds of patients, to collect safety and tolerability data and to test for histological efficacy. These trials are exploratory, meaning that all histological changes need to be documented and studied including; changes in steatosis, hepatocyte ballooning and inflammation, fibrosis, NAFLD activity score (NAS), area of fibrosis (i.e. micromorphometry, collagen proportionate area), activation of fibrogenic cells (alpha smooth muscle actin staining) etc. Not all of these lesions are necessarily associated with clinical outcomes and therefore most are not considered approvable outcomes from a regulatory perspective. For instance the NAS has not been shown to clearly predict outcomes and there are conflicting data as to whether baseline values or longitudinal changes are associated with fibrosis progression, which may be because NAS is an aggregate score of both disease activity and steatosis. Regardless, what is important in these trials is an exhaustive description of histological changes and particularly of all elementary lesions; this will help identify whether there is potential for histological benefit. However, at this stage, it is also crucial to include the two composite histological outcomes that are reasonably likely surrogates for conditional approval in registrational trials (discussed later). Moreover, the population of patients included in these trials needs to be very similar to the one agreed upon for inclusion in phase III, registrational trials. This is because these phase IIb trials will directly influence the design of subsequent trials. It remains to be seen if non-invasive biomarkers will replace histology in phase IIb trials in the near future. For this to happen, it will be necessary to demonstrate the diagnostic value of these biomarkers for histological changes, and to develop an approved qualification for use. For the moment, only measurement of...
elastometry by Fibroscan is approved for detecting stiffness of inner organs, although it is not yet approved to detect changes in liver stiffness induced by therapeutic interventions.

**Overall objectives of phase III trials**

Phase III, or registrational trials, are intended for marketing application. Because of the unmet need in NASH therapy, the regulatory authorities have now agreed upon a two-step process designed to accelerate drug approval: an early, conditional approval, based on achieving reasonably likely surrogates and a subsequent, definitive approval, based on achieving a generally accepted surrogate or hard clinical outcomes (for a detailed review see). These are usually combined in a single, long-term trial, including an interim analysis after 12 to 18 months of therapy, followed by an outcome trial lasting for several years. There are two reasonably likely surrogates for conditional approval: i) resolution of NASH without worsening of fibrosis (i.e. a numerical increase in fibrosis stage); ii) a one or more stage reduction in fibrosis without worsening of NASH (i.e. a numerical increase in the ballooning or inflammatory grade).

Whether one or other of these composite endpoints can be met after a 12 to 18-month treatment period ultimately depends on the mechanism of action of the drug. A drug with a dominant “antifibrotic” activity will more likely meet the fibrotic endpoint, whereas a metabolic modulator or a drug controlling liver cell injury and inflammation will more likely be successful on the steatohepatitis endpoint, even if fibrosis may improve subsequently. Of course, the ultimate histological benefit cannot always be anticipated based on the mechanism of action in preclinical models or early human studies and some drugs may have pleomorphic effects. Nevertheless, what ultimately matters is whether the drug candidate inhibits the progression to cirrhosis and results in a reduction in hepatic clinical events (complications of cirrhosis), liver transplantation or death. These are, precisely, the endpoints for the outcome trial; if the benefit over placebo is considered substantial, the drug candidate may then obtain definitive approval.

Hepatocellular carcinoma is a major complication of NASH, but current trials do not specifically address a reduction in the incidence of hepatocellular carcinoma as a main outcome. Instead, the number of incident tumours are part of the collated total of hepatic clinical events, which includes other complications of cirrhosis in the outcome trials. The main reason for this is that hepatocellular carcinoma can arise throughout the whole spectrum of NAFLD (steatosis, steatohepatitis with or without fibrosis, with mild or severe activity and cirrhosis). Since phase III trials only include patients from part of this spectrum, a reduction in incident hepatocellular carcinoma would not truly reflect an overall clinical benefit in terms of hepatic carcinogenic complications. For the same reason it is unknown if resolution of steatohepatitis, through a reduction in ongoing cell injury and inflammation, would result in a reduced rate of hepatocellular carcinoma. Animal models have identified numerous carcinogenic mechanisms that may not be primarily driven by steatohepatitis including: a reduction in hepatocyte apoptosis in the steatotic liver of obese rodents, regardless of the presence of fibrosis; an increase in circulating levels of insulin-like growth factor-1; increased lipogenesis contributing to liver oncogenesis; a tumour-promoting effect of dietary obesity through low grade systemic inflammation; frequent chromosomal alterations in NAFLD hepatocellular carcinoma, to name just a few. Therefore, it is unclear whether the mere reduction in fibrosis stage or the resolution of NASH will have a meaningful effect on the rate of hepatocellular carcinoma.

**A review of the histological surrogates and their validity as predictors of cirrhosis occurrence**

**Changes in fibrosis stage**

In most chronic liver diseases, patients die of complications of cirrhosis, including primary liver cancer. However, cirrhosis is the result of a protracted fibrogenic process, which is artificially segmented into fibrosis stages, based on landmarks defined by pathology. These stages primarily reflect changes in lobular architecture and not the amount of fibrosis. Regardless, patients need to travel through this fibrogenic process, a journey that is lengthy and to a certain extent reversible. Therefore, it does make sense to try and track the progression to cirrhosis by measuring changes in fibrotic stages along the way, both for natural history studies and for drug trials. An efficient drug would either delay or reverse the fibrogenic process, which could be captured by measuring changes in fibrosis stages. At this point, it is important to keep in mind that the fibrosis stage is only a surrogate for cirrhosis and its complications. Contrary to cirrhosis, stage 1, 2 or 3 fibrosis does not cause a patient’s death. The question is, how can we define changes in fibrosis in such a way that they act as a robust surrogate for the occurrence of cirrhosis?

Whether one considers the histological definition of fibrosis stages as arbitrary or not, there are clear shortcomings. Firstly, the one to four stage division does not reflect a linear increase in the amount of extracellular matrix (mostly collagen). Micromorphometric studies of the fibrosis area have shown that there is relatively little difference in the amount of fibrosis between stages 0, 1 and 2 compared to bridging fibrosis, particularly advanced bridging (METAVIR F3). There is then a major increase at the cirrhotic stage. This lack of linearity is only a problem when trying
to measure the potency of an antifibrotic drug by a one stage reduction, regardless of which one stage it is. The second shortcoming is that despite their sequential labelling, it is not yet established that progression to bridging fibrosis for instance, necessarily follows the sequential pattern, from stage 1 to 2 then 3. These two shortcomings raise substantial difficulties when using the current fibrosis staging system to study the antifibrotic effect of a drug and its ability to prevent fibrosis progression.

However, fibrosis stages are defined, and keeping in mind that they are no more than surrogates, natural history studies can bring validity or clinical relevance to their definition. The association with clinical outcomes would increase the value, as a surrogate, of carefully defined changes in fibrosis stage. Accumulating data in several chronic liver diseases, including NASH, has shown an increase in liver-related events (i.e. complications of cirrhosis, hospitalisation for decompensated cirrhosis, liver transplantation, liver-related death) in patients with bridging fibrosis compared to the absence of fibrosis.\(^7\) This increase is of course much higher at the cirrhotic stage. For now, the data on bridging fibrosis and cirrhosis as predictors of liver-related events are the most robust we have. This turning point in the natural history of chronic, fibrotic liver diseases has provided the rationale used to define the treatment population; before the era of highly potent, well tolerated antivirals, the indication for therapy in patients with HCV was F2 METAVIR (bridging fibrosis). In NASH, a few studies with a higher number of patients and a longer follow-up became available and have shown a significant increase in liver-related events, even before the bridging fibrosis stage, at NASH Clinical Research Network (CRN) stage 2.\(^7\) It is still controversial whether this also applies to stage 1,\(^7\) as one study has shown,\(^7\) but ultimately the demonstration may simply be dependent on following a larger cohort over a longer period of time. Nevertheless, all these studies come from tertiary referral centres and retrospectively analyse a highly selected population that underwent liver biopsy and survived competing mortality. While a large body of data demonstrating an increase in liver-related events at all fibrosis stages will provide a rationale for deciding which patients require pharmacotherapy, it does not necessarily mean that a one stage reduction is a valid surrogate. For example, if for an early fibrosis stage the increase in the hazard ratio for liver-related events is only marginal, albeit statistically significant, this may not sufficiently equate to a meaningful reduction in the risk of progression to cirrhosis in patients at that early stage.

**Key point**

Only the proportion of patients whose fibrosis stage worsens will directly impact on the number of patients progressing to cirrhosis, making it unlikely that improvement of fibrosis can act as a valid surrogate of progression to cirrhosis.

The likely surrogate that best predicts progression to cirrhosis has yet to be defined in such a way that it is achievable within a one to two-year time frame. Longer-term exposure may directly capture progression to cirrhosis with a sufficient number of events, but this is not compatible with the shorter time frame of the accelerated conditional approval. A more stringent endpoint was proposed, such as a two or more stage reduction in fibrosis. This removes some of the uncertainty around the non-linearity of the fibrotic deposition by providing a more robust reduction in fibrosis, but it still does not directly measure deterioration. The same could be said for the complete disappearance of fibrosis, with the added difficulty of low rates of response making even statistically significant differences unconvincing for regulators. Given that bridging fibrosis and cirrhosis are both associated with liver-related mortality in all studies to date,\(^7\) a more robust way to define improvement or deterioration in fibrosis would be the proportion of patients who no longer have bridging fibrosis (if they had it at baseline) or the proportion of patients that progressed to bridging fibrosis (or to cirrhosis if they started at the bridging stage). However, the current NASH CRN classification

**Limitations of the current definition of the antifibrotic response**

Currently, the consensus among regulatory bodies and experts is that a one or more stage reduction in fibrosis without worsening of steatohepatitis is a likely surrogate for preventing progression to cirrhosis, and it is actually the basis for conditional approval\(^6\) (although it has not been used as such at the time of writing). Fibrosis reversal is a legitimate aim when dealing with a progressive fibrotic disease and many studies have shown major fibrosis reversal with therapies that eradicate the cause of liver injury (antivirals in viral hepatitis, immunosuppressants in auto immune hepatitis, etc.). However, there is no likely curative treatment for NASH on the horizon, but rather drugs that control or slow down the progression of the disease in a minority of responders, without eradicating the “cause”. Consequently, only the proportion of patients whose fibrosis stage worsens, not the proportion of those that improve, will have a direct impact on the number progressing to cirrhosis. While in the best case scenario the two can go together, evolving in opposite directions, it is not always the case; a recent therapeutic trial has shown more patients improving fibrosis stage on active drug than on placebo, but the same proportion with worsening of fibrosis in the two arms.\(^9\) Obviously, patients that experience an increase in fibrosis stage, not those with a stage reduction, are the ones that progress to cirrhosis. Therefore, it is doubtful that an improvement in fibrosis is a valid surrogate of progression to cirrhosis, unless an unrealistic 100% response rate is achieved. Other theoretical examples where the current definition of antifibrotic response either cannot predict the progression to cirrhosis or cannot be unambiguously interpreted are shown (Fig. 1).
Antifibrotic activity. Despite a higher rate of fibrosis regression with Drug A than with Drug B, a similar proportion of patients with a one or more stage worsening would be expected to result in a similar proportion of progression to cirrhosis; (B) Despite a similar proportion of patients with worsening of fibrosis (which is expected to drive progression to cirrhosis) the much higher proportion of patients with fibrosis improvement makes it hard to conclude against an overall antifibrotic benefit; (C) A higher proportion of both improvement and worsening of fibrosis with Drug A vs. Drug B makes it uncertain which of A or B has the best antifibrotic activity.

Antifibrotic endpoints other than histological stage changes
Micromorphometry correlates with Ishak fibrosis stage and with portal hypertension. Because it is a quantitative variable, it could be more sensitive to change than the histological stage, and therefore could provide additional information on the antifibrotic potency of a drug. The area of fibrosis has been shown to correlate with clinical outcomes in several studies, a pre-requisite for it becoming a likely surrogate. Changes in the area of fibrosis have been shown to predict clinical decompensation in patients with HCV recurrence after liver transplantation, and in patients with NASH and advanced fibrosis. The measurement of the area of fibrosis by digital image analysis has been used in several antifibrotic trials, but it is not yet an acceptable endpoint, only a secondary endpoint intended to support evidence for an antifibrotic effect.

Resolution of steatohepatitis
One of the very early findings about the natural history of NAFLD was that steatohepatitis bears a burden of hepatic morbidity and mortality that steatosis does not. Age and sex-standardised mortality ratios were higher for steatohepatitis than for the general population, while this was not the case for steatosis. Besides, it is believed that steatosis has little or no fibrogenic potential, while almost all cases of fibrosis develop in patients with steatohepatitis. It is therefore logical that the disappearance (resolution or reversal) of steatohepatitis would be beneficial as it would place the patient in a very low risk category for disease progression.

Accumulating evidence has strengthened and refined this concept. Patients with steatohepatitis are older and have higher levels of insulin resistance than those with steatosis. They also have a higher prevalence of complications of the metabolic syndrome and a more severe phenotype. Their hepatic disease is more advanced, with higher overall levels of fibrosis, more marked biochemical injury (increased ALT), and, by definition, more necroinflammatory activity. Thus, steatohepatitis represents a more advanced form of NAFLD than steatosis and is associated with more advanced metabolic disease. Most importantly, studies with serial liver biopsy have shown that steatohepatitis, mainly through chronic, uncontrolled necroinflammation, promotes fibrogenesis. Earlier data indicated that necroinflammation is the main risk factor associated with fibrosis progression and this has recently been confirmed in a study comparing fibrosis progression in patients with steatohepatitis vs. those with steatosis only. Also, on follow-up biopsies, the occurrence of advanced fibrosis coexists with increasing grades of necroinflammation (i.e., disease activity). Progression from steatosis to NASH increases the risk of significant fibrosis (stage 2 or higher) 7.2-fold on follow-up liver biopsy. Conversely, when patients progress from steatosis (without NASH or fibrosis) to significant fibrosis, they almost always develop steatohepatitis in the process. Moreover, there seems to be a quantitative relationship between disease activity and fibrosis deposition; both baseline necroinflammatory scores and increases in those scores are associated with more fibrosis in studies with serial biopsies. Ballooning grades are correlated with clinical hepatic outcomes in univariate analyses. Finally, interventions that modulate disease activity impact on the rate of fibrosis. Pharmacological treatment-induced changes in disease activity (the sum of inflammation and ballooning) are positively correlated with changes in fibrosis; a reduction in disease activity is associated with reduced fibrosis, whilst an increase in activity is associated with increased fibrosis.

Therefore, the data mentioned above is good clinical evidence that steatohepatitis is a major factor behind fibrosis build-up and consequently...
disease progression in NASH. Current phase III trials test the hypothesis that removing this driving force will result in less progression towards cirrhosis. If proven, this will validate its value as a reasonably likely surrogate. It will also, conceptually, promote the idea that fibrosis reversal can be obtained either through a direct antifibrotic action or, indirectly, as a consequence of removing the cause of fibrosis by treating NASH. Of note, the objective is not an improvement in NASH activity but rather its disappearance, a more stringent requirement. The definition of NASH resolution is now consensual among experts (Liver Forum, Manuscript in preparation); it requires disappearance of hepatocyte ballooning and either disappearance or persistence of minimal, residual, lobular inflammation. It is now being used as a regulatory outcome in several large, international, phase III studies91,92 and has been reported on in at least two recent therapeutic trials.39,40 Unfortunately, direct comparison with response rates from older studies is not possible as the definition of resolution of NASH was either different or insufficiently detailed.12,93

Conclusion
Many endpoints are currently being measured in NASH trials in an attempt to find the best predictors of a drug candidate’s biological activity and efficacy. One of the major breakthroughs in the field was the construction of a regulatory approval framework whereby endpoints that can be achieved within a reasonably short timeframe, compatible with clinical trials, were granted reasonably likely surrogate status. The second major breakthrough was the recognition that preventing progression to cirrhosis should be a major therapeutic objective, on a par with the documentation of liver-related events such as cirrhosis complications, liver transplantation and death. The major challenge of ongoing registrational trials will be to demonstrate that achieving reasonably likely surrogates, such as resolution of NASH and an improvement in fibrosis stage, does indeed translate into a reduction in the rate of progression to cirrhosis. While waiting for regulatory acceptable biomarkers of fibrosis, a better histological definition of antifibrotic activity that considers both stage improvement and stage worsening is critical.

Conflict of interest
Professor Ratziu reports consulting fees from Allergan, Galmed, Genfit, Intercept, Pfizer, Boehringer-Ingelheim, Novo-Nordisk, Grants from Intercept, Gilead, outside the submitted work.

Please refer to the accompanying ICMJE disclosure forms for further details.

Acknowledgements
I wish to thank Professors Hannele Yki-Jarvinen and Jacob George for insightful discussions and help with the manuscript.

References

Author names in bold designate shared co-first authorship
