EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

15 November 2018
EMA/CHMP/299976/2018
Committee for Medicinal Products for Human Use (CHMP)

Reflection paper on regulatory requirements for the development of medicinal products for chronic non-infectious liver diseases (PBC, PSC, NASH).

Draft

Draft agreed by Gastroenterology Drafting Group	October 2018
Adopted by CHMP for release for consultation	15 November 2018
Start of public consultation	16 November 2018
End of consultation (deadline for comments)	31 August 2019

Comments should be provided using this template. The completed comments form should be sent to GastroenterologyDG@ema.europa.eu

Keywords	Non-alcoholic steatohepatitis (NASH), Primary Biliary Cholangitis (PBC), Primary Sclerosing Cholangitis (PSC), clinical trial design, endpoints, conditional marketing authorisation, unmet medical need, liver biopsy

30 Churchill Place ● Canary Wharf ● London E14 5EU ● United Kingdom
Telephone +44 (0)20 3660 6000 **Facsimile** +44 (0)20 3660 5555
Send a question via our website www.ema.europa.eu/contact

Reflection paper on regulatory requirements for the development of medicinal products for chronic non-infectious liver diseases (PBC, PSC, NASH).

Table of contents

1. Introduction

Chronic, non-infectious liver diseases are a medical field of high unmet medical needs. At the same time, the specifics of the diseases create major challenges for the development of new medicinal products. This reflection paper restricts the current regulatory approach to 3 different disease entities primary biliary cholangitis (PBC), primary sclerosing cholangitis (PSC), and non-alcoholic steatohepatitis (NASH) for which recent efforts are undertaken to bring new medicinal products to the market.
It is anticipated that many of the problems raised and potential solutions described in this reflection paper, may be transferrable to other chronic liver diseases.

2. Scope

As a reflection paper, this guidance document provides a high level description of the requirements for drug development in the field. For all three disease entities dealt with in the paper, the regulatory experience with the licensing of new medicinal product is limited. Therefore, this paper aims at a preliminary definition of development strategies only, which, in the case of several successful MAAs occurring in the future, will have to be refined, and may finally be superseded by full guidance documents.

3. Legal basis and relevant guidelines

This document should be read in conjunction with the introduction and general principles and part I and II of the Annex I to Directive 2001/83/EC as amended. Applicants should also refer to other relevant EU and ICH guidelines (in their current version) and regulations, especially the following:

- Reflection paper on assessment of cardiovascular safety profile of medicinal products (EMA/CHMP/50549/2015)
- Reflection paper on the use of extrapolation in the development of medicines for paediatrics. (EMA/189724/2018)
- Guideline on clinical development of fixed combination medicinal products. (EMA/CHMP/158268/2017)
- ICH E9(R1) Draft Addendum on estimands and Sensitivity Analysis in Clinical Trials (EMA/CHMP/ICH/436221/2017)

4. Recommendations

4.1. General considerations

Chronic liver disease is a slowly developing process, and many patients do not develop relevant disease sequelae, and/or symptoms over even a considerable time of observation, and the development of end-stage liver disease may be a process of years, if not decades. All three diseases under consideration will be difficult to be studied for long-term outcomes over a reasonable time span (the term “long-term outcome” is used in the following for events such as liver transplantation and death, as well as clinical events of decompensation of liver cirrhosis which are otherwise also termed “hard outcomes”).

An acceptable regulatory strategy for companies developing new agents in the disease area, may be to look for intermediate endpoints for which a reasonable assumption for the prediction of long-term outcomes can be made. These reasonable assumptions are usually based on associations with regard to risk factors for the long-term outcomes in observational natural history cohorts and the biological plausibility attributed (the term "intermediate endpoint" will be used throughout in the following for events otherwise also termed "interim" or "surrogate" endpoint).

Strictly speaking, however, such endpoints are not validated in the sense that positive changes for the surrogate as well as the long-term outcome have repeatedly and consistently been demonstrated for therapeutics. Due to the largely unmet medical need in the field, a strategy to obtain an early approval of new compounds based on these intermediate endpoints, however, could be considered. This strategy will require the confirmation of efficacy (and safety) of the compound after approval (including availability on the market) documenting the effects on long-term outcomes. Such a strategy could be acceptable as long as an unmet medical need can still be reasonably concluded[1] If such a strategy is intended, however, the evidence at the time of evaluation of the intermediate endpoints has to be such that it allows the conclusion of a positive risk-benefit ratio independent from the presence of an unmet medical need.

In the situation of unmet medical need, the use of placebo as comparator would be the only acceptable way to demonstrate efficacy. However, the authorisation of new substances may trigger a revision of the acceptance of placebo as comparator in the future.

The acceptance of the mentioned regulatory strategy has to be regarded to be a case by case issue.

For the three disease entities, this document will display endpoints which can currently be considered acceptable surrogates for the manifestation of end-stage liver disease (the intermediate endpoints), as well as those deemed suitable for the confirmation of these surrogates (the long-term endpoints). The specifics to confirm (and thus validate) these surrogates will be dealt with in the relevant chapters.

These intermediate endpoints (as well as the long-term endpoints) are currently partly or mainly based on the histological evaluation of liver biopsies. Liver biopsy and histology have been widely criticized for sampling error and intra- and inter-observer variability[2]. However, potential non-invasive methods do currently have insufficient, and especially insufficient disease specific, validation data available, and therefore histology is in most cases still regarded to be the state of the art for the diagnosis, and especially for the follow-up of the course of the diseases, in particular for the purpose of clinical trials.

Liver biopsy, however, is also unwanted due to its patient burden, invasiveness, and the associated risks of morbidity[3] and potentially even mortality.

Therefore, this reflection paper also calls for the further development of non-invasive methods to replace liver histology in the future, be it serological markers, or imaging methods. It is therefore recommended that future applicants should use development programmes aimed at producing evidence for the approval of new medicinal products also for the further validation of novel methods, intending to replace histology in the future.

Possible targets of estimation that define treatment effects of interest in the three disease entities are also considered.

4.2. Non-alcoholic steatohepatitis

4.2.1. Short characterisation of the disease

Non-alcoholic steatohepatitis (NASH) is considered the progressive phenotype of non-alcoholic fatty liver disease (NAFLD), which itself is the most prevalent chronic liver disease worldwide with an estimated prevalence in the Western world of around 25% [4][5], and it is estimated that about 20% of these suffer from NASH[6]. The progression is related to the development of liver cell stress, subsequent inflammation, and fibrosis with the potential development of cirrhosis, and end-stage liver disease. NASH is also a relevant risk factor for the occurrence of hepatocellular carcinoma. NASH associated cirrhosis/end-stage liver disease is expected to represent the highest share of patients referred for liver transplantation in the future[7]. NAFLD as well as NASH are associated with other comorbidities and risk factors such as obesity, arterial hypertension, diabetes mellitus type 2 (T2DM), atherogenic dyslipidaemia, and others. The disease – although genetic factors have also been identified [8] – is thought to be the consequence of hyperalimentation, and has been regarded to be the hepatic manifestation of the so-called metabolic syndrome[9].

From a diagnostic point of view, the diagnosis of NASH is one of exclusion (involving the exclusion of relevant alcohol intake, and infectious and non-infectious liver disease) as well as positive confirmation of the features by liver biopsy and histology, the latter relating to the pathognomonic features of steatosis, hepatocellular ballooning, lobular inflammation, and fibrosis[10]. Although health-related quality of life may be impaired [11], symptoms do not play a relevant role in (non-cirrhotic) NASH. Also the awareness with regard to the disease and of the associated risks is poor[12].

The natural history of NASH has not been fully elucidated, and further efforts are needed to clarify important aspects, e.g. overlap of progression and regression. The risk of progression to end-stage liver disease is largely related to the baseline fibrosis grade[13]. The progression of fibrosis is estimated to be slow, and progression of 1 fibrosis stage is estimated to occur at a mean of more than 7 years (7.7 years; 95% CI 5.5-14.8 y) [14][15][16].

4.2.2. Selection of patient populations

The usual principles of the selection of study population, such as being representative of the target population, and being well balanced with regard to demographic characteristics and co-morbidities are of course applicable to NASH. The more specific requirements are dealt with in the following:

As mentioned above, the diagnosis of non-alcoholic steatohepatitis is a diagnosis of mixed exclusion of other relevant diseases, as well as a positive diagnosis which is mainly reliant on liver biopsy with histology. Histology is currently considered to be the gold standard for finally securing the diagnosis, as well as determining the severity of disease, and is also recommended as part of clinical practice. A selection of patients on the basis of symptoms is usually not possible, and the (long-standing) presence of features of the metabolic syndrome can only be used as a trigger to identify potential study participants.

The risk of progression to end-stage liver disease, liver transplantation and death has been demonstrated to be independently associated with the stage of liver fibrosis only, with only minimally increased risk for stage 1 patients [17]. Fibrosis stage 1 patients are therefore currently only recommended for inclusion in therapeutic trials in NASH for exploratory purposes.

Therefore, the "natural" selection of patients with an unmet need for treatment in NASH relates to patients with (fibrosis) stages 2-4 NASH.

Inclusion of patients in fibrosis stages 2 and 3 should additionally be based on the disease activity / grading because developments of regression and progression may overlap, and a (albeit univariate only) risk of progression has also been associated with higher degrees of ballooning and inflammation. The patient population should be included based on a valid grading system for NASH with minimal requirements for the presence of cell stress (ballooning), as well as inflammation (lobular inflammation). The NASH-CRN (Non-alcoholic steatohepatitis clinical research network) histological scoring system currently appears to be the best validated and most widely accepted system [18]. A total NAS (NAFLD activity score) of at least 5 appears acceptable but a score of 4 can be accepted as well, if it is not based on a high contribution of the steatosis grade alone and minimal requirements for relevant ballooning and lobular inflammation are fulfilled (scoring of at least 1 in each of these 2 components). Although the NASH-CRN grading system is the recommended grading system, patients may also be included based on potentially other grading systems for NASH, provided the validation of respective grading systems is substantiated.

In patients with manifest cirrhosis (=fibrosis stage 4), the presence of such a rigorous minimal grade is less critical, because the risk of (clinical) progression is thought to be high based on the presence of cirrhosis alone. Nevertheless, in so-called burnt-out NASH cirrhosis or patients initially diagnosed with cryptogenic cirrhosis[19], if definite NASH is not present, all of the following should be available in order to make the diagnosis NASH sufficiently likely: historical biopsies with presence of unequivocal NASH, a high likelihood of NASH based on non-invasive testing (biomarker and imaging), and presence of associated co-morbidity (e.g. obesity with T2DM).

Patients with decompensated cirrhosis represent a particularly vulnerable subset of patients. A relevant amount of mechanistic, as well as clinical efficacy and safety data on an investigational compound may be required before the inclusion of such patients into clinical trials. Due to the fact of increased risks of biopsies in this population, historical biopsies (with presence of cirrhosis) together with symptoms of decompensation may be used as inclusion criteria in this population.

The multi-stakeholder composed Liver Forum has recommended that histology should always be available, also in early clinical trials [2021], and inclusion of patients should always be based on histological evaluation (grading and staging). Deviations for exploratory clinical trials, e.g. using imaging methods, or biomarkers (or a combination of those) only, are possible if based on sound scientific principles, for which the uncertainties can be quantified and later stage trials be planned accordingly. [22]

The positive influence of weight reduction on NASH has clearly been demonstrated[23]. Therefore, before inclusion of respective patients into clinical trials for NASH, it is recommended that patients should have undertaken at least one unsuccessful attempt with weight-reducing diet. Co-morbidities, such as e.g. diabetes, hypercholesterolaemia, and hypertension should adequately and stably be treated at the time of inclusion.

Important factors to be considered in all populations are the presence of co-morbidities (e.g. diabetes), and stratification for these factors could be advisable to allow a balanced evaluation of these covariates.

In summary, for the purpose of therapeutic clinical trials, NASH may be considered in three broad categories:

a. Definite NASH based on histology with demonstration of NAS≥5 (or NAS ≥4 with all components of at least 1) and fibrosis stage 2-3

b. Compensated NASH-cirrhosis based on histology with fibrosis stage 4 and NASH diagnosis based on either NAS>5 (or NAS ≥4 with all components of at least 1) or the availability of historical

histology proving NASH, non-invasive tests pointing to NASH (serological markers, imaging), and relevant co-morbidity risk-factors (obesity and type 2 diabetes mellitus (DM))

c. Decompensated NASH Cirrhosis: Presence of historical biopsy data showing unequivocal NASH as well as cirrhosis; symptoms of decompensation.

4.2.3. Study design and endpoints

The natural history of NASH is assumed to end with the manifestation of cirrhosis in the liver, and the subsequent development of portal hypertension and its sequelae, and decompensation of liver function, which ultimately results in liver associated death, or liver transplantation. Because NASH is also associated with a multitude of risk factors for cardiovascular disease (hypertension, obesity, atherogenic dyslipidaemia, and type 2 diabetes), a relevant proportion of patients will also be prone to causes of death other than liver related ones, mainly cardiovascular.

The "natural" long-term endpoint in clinical trials for NASH would therefore be the combination of all-cause mortality, liver transplantation, and the manifestation of decompensation (variceal bleeding, ascites, encephalopathy etc.).

Stage 2 and 3 fibrosis:

The time to manifestation of long-term outcomes is currently largely unknown, and reasonably sized trials in patients with the earlier stages of disease (such as fibrosis stage 2 and 3) with the primary aim to demonstrate an effect on survival free of liver transplant and decompensation events might be unfeasible. Therefore, efficacy endpoints reflecting a substantial increase in the risk of disease progression (to the events described) are needed. The histological diagnosis of cirrhosis has been proposed to represent such an endpoint, and is regarded to be an acceptable surrogate and can therefore be part of the long-term endpoints. Similar arguments have been accepted for a model for end-stage liver disease (MELD) score above the threshold of 14. The long-term outcome for the demonstration of efficacy in NASH is therefore proposed to be a composite endpoint with the components all-cause death, decompensation of liver disease (with a complete listing), as well as (histological) diagnosis of liver cirrhosis and MELD>14

However, due to feasibility issues in long term studies and the unmet medical need in NASH, an earlier evaluation of efficacy, with an overall shorter duration of clinical trials is warranted and intermediate endpoints reasonably predicting the long-term outcome have been advocated[24].

Acceptable intermediate endpoints would consist of two composite endpoints to be evaluated at the individual patient level:

1. The resolution of NASH – with the presence of any grade of steatosis, no ballooning, and only minimal (grade 1) lobular inflammation and – at the same time – no worsening of the stage of fibrosis.

2. The improvement of fibrosis by at least 1 stage without any worsening of NASH (no worsening of ballooning and lobular inflammation, a 1 grade change in steatosis may be acceptable).

Efficacy in these two composites should be demonstrated in co-primary fashion, meaning that both will have to independently demonstrate a statistically significant and clinically relevant difference to placebo. This requirement is thought to take account of the uncertainties associated with a strategy to account for the long-term outcomes later.

Stage 4 fibrosis (NASH cirrhosis)

In liver disease where cirrhosis has already been manifested, the use of the above mentioned long-term composite is not possible. An acceptable endpoint for patients with already existing cirrhotic liver disease at inclusion would therefore consist of the composite of all-cause death and liver decompensation events. However, because liver cirrhosis represents a wide spectrum of disease, it is currently unclear whether such an endpoint is feasible. When the intention is to use this long-term endpoint in the cirrhotic population, the study population should be enriched with patients with advanced cirrhosis.

In case the need to use intermediate endpoints in this population is identified, a reasonable endpoint for the general non-decompensated population, could intuitively be the reversal of cirrhosis (e.g. defined as "improvement of liver cirrhosis to non-cirrhotic liver disease (1 or more point improvement in fibrosis stage"). At this point of time, however, the data available to demonstrate that reversed cirrhosis does indeed also reverse or influence the final prognosis substantially, is considerably less profound than the association shown for progressing disease. Such a trial would therefore need the substantiation of the claim that the prognosis of reversed cirrhosis is similar to the prognosis of (untreated) earlier stages of fibrosis in progressive disease (e.g. from other disease areas such as chronic infectious liver disease; ie Hepatitis C or B). Moreover, this endpoint should be appropriately backed by additional, secondary outcomes, based on non-invasive markers of disease (imaging techniques, determination of liver stiffness, biomarkers) as well as the available (descriptive) data on decompensation events, liver transplantation, and death.

In a situation when relevant proportions of patients with advanced cirrhosis are included, an acceptable endpoint would be the occurrence of decompensation events since the prognosis for patients with decompensated cirrhosis is markedly worse than those with compensated cirrhosis[25].

Other potential endpoints (e.g. lowering of MELD score below a certain threshold, or of the HVPG below 10 mm Hg[26]) are also possible based on specific justifications.

The need for addition of post-marketing observations with regard to the manifestation of end-stage liver disease and death (=the long-term outcome observation) will in these cases be assessed based on the overall substantiation of the clinical usefulness of the primary endpoint used and the data on the secondary outcomes.

In the special group of decompensated cirrhosis, a therapeutic effect should be demonstrated based on the endpoint all-cause mortality/survival. Liver related death, and liver-related death/ transplantation could be supportive endpoints.

Additional considerations on mode of action

As a simplified pathophysiology of NASH, it has to be assumed that the liver cell toxicity caused by the overload in fat causes inflammation, which itself is the final trigger of fibrosis development. Therefore, it has been assumed that the appropriate target of medicinal products would be mechanisms preventing fat toxicity and/or decreasing inflammatory activity, which would finally lead to beneficial effects in fibrosis. However, new substances primarily targeting the development of fibrosis are currently under development, and it is therefore considered important to reflect whether a decrease in fibrosis stage without any or only minor influence on the fat accumulation in the liver, liver cell stress (ballooning) and inflammation could be appropriate as treatments and benefit patients in the long term. This is considered an uncritical question as long as long-term endpoints are used as objectives in clinical trials. However, in case an intermediate endpoint strategy is followed, the above mentioned two composite endpoints may be impossible to be used due to the fact that a resolution of NASH endpoint is not within reach of such compounds. If an intermediate endpoint strategy is used in such

compounds, it is currently recommended to use a stronger endpoint denoted as a composite at the individual patient level such as "fibrosis regression of at least 2 stages without worsening of NASH", in which stage 2 fibrosis patients would need to achieve complete resolution of fibrosis, and patients with stage 3 would need to regress into a stage associated with only minimally increased risk for disease progression ("no worsening of NASH" could be defined similar to the above).

This requirement similarly applies to patients with cirrhosis. Although it might be possible to show that the reversal of cirrhosis benefits patients with other liver diseases (e.g. with data from Hepatitis C or B trials) in the long-term, there remain important questions with regard to the ongoing primary insult (the fat associated necro-inflammation), which cannot be solved with the data from infectious liver disease, because these have – in their vast majority – been derived from patients with sustained viral response, and thus an almost complete suppression of the inflammatory insult.

Duration of trials

The currently published phase 2 data for substances under development have mostly evaluated parts of the above proposed endpoints only. Therefore, uncertainty exists with regard to the duration of trials, both in terms of the time needed for interim evaluation with the intermediate endpoints, as well as for the time needed to show relevant effects on the long-term composite endpoint. As a general rule, a two-year interim evaluation, and a 5-year final evaluation may be considered appropriate. However, this can be modified with factors like size of the trial, activity of the investigational compound, patient characteristics, and the requirements with regard to statistical rigor. The final evaluation would be expected to be usually planned with an event-driven evaluation, and therefore, a fixed duration may not be appropriate to be planned with.

Target of estimation (estimand)

The scientific question(s) of interest, i.e., what the trial seeks to address and ultimately, the target of estimation (estimand) should be specified. The trial planning, design, conduct, analysis and interpretation must be aligned with the estimand. It is referred to ICH E9(R1) Draft Addendum on estimands and Sensitivity Analysis in Clinical Trials (EMA/CHMP/ICH/436221/2017).

In order to determine the appropriate strategy for a trial in NASH, a full review of potential intercurrent events is necessary. Relevant intercurrent events expected are those associated with almost all clinical trials, such as treatment discontinuation and use of additional medication. Contrary to other fields of development, the use of rescue medication may – for the time being – not be relevant because no specific treatments are available, but could become more relevant in the future. However, a change in background medication (including excessive life-style changes with weight loss, or uptake of relevant alcohol intake) may relevantly affect the outcome, and may need to be considered.

For the intermediate endpoints, the outcome regardless of the occurrence of intercurrent events is of primary interest (i.e. a treatment policy strategy discussed in the addendum).Therefore, data with regard to the outcomes of interest should be collected independently from the occurrence of an intercurrent event. Data that is nevertheless not collected, for example in case the endpoint is based on liver biopsy and the biopsy is missing or not evaluable, results in a missing data problem with regard to subsequent statistical inference.

Choices made regarding statistical analysis, including the handling of missing data, must be aligned with the target of estimation. Considering a patient with missing data as a non-responder usually results in a conservative estimate of the treatment effect with regard to the question of primary interest, but alternative handling of missing data may also be acceptable (possibly taking occurrence of intercurrent events and the reason for missing data into account). For example, for patients on treatment who refuse biopsy, replacing missing data using multiple imputation based on response

probability of patients still on treatment (possibly taking additional covariates into account) could be considered.

The outcome regardless of occurrence of intercurrent events is also of primary interest for the hard endpoints. Aiming at a complete follow-up for the outcome events is of particular importance as patients that are not completely followed are likely to have a different prognosis than patients who complete the study, implying that censoring such patients is probably informative and leads to bias. As a biopsy during the follow-up is only scheduled if there is a high likelihood of a cirrhosis (e.g. based on surveillance with non-invasive methods such as fibroscan), non-performance of a scheduled biopsy should be considered as an event.

Combination treatment

It has been advocated, based on the results of currently available phase 2 trials, that a satisfactory treatment of NASH might only be possible, if new investigational compounds are combined, ideally with a combination of two different principles of action, such as e.g. anti-fibrotic, and anti-inflammatory [27]. Whereas such a strategy can be followed from a theoretical point of view, potential applicants should move forward carefully with such development programmes in a situation with no established therapies available.

The main considerations of the Guideline on clinical development of fixed combination medicinal products (EMA/CHMP/158268/2017) will have to be taken into account when embarking on a co-development, and ultimately on a fixed-dose combination product.

The expectations from the regulatory side would be that the combination is based on valid therapeutic principles, but also that for each of the substances involved, the contribution to the therapeutic effect is demonstrated.

It will also be expected that the properties of the single substances are fully explored and described either before or during the development of the combination treatment. Also, referring mainly to other disease areas, it will be expected that either a second line treatment is investigated, which has to include the establishment of a definition of an insufficient response to a standard treatment (or at least one of the combination partners), or – in case an initial combination treatment is aimed at – the definition of a patient group with a very high risk of progression.

4.3. Primary biliary cholangitis

4.3.1. Short characterisation of the disease

Primary biliary cholangitis (PBC), previously known as primary biliary cirrhosis[28], is a chronic, slowly-progressive autoimmune cholestatic liver disease[29]. The disease is mainly diagnosed in female patients with a ratio of about 10:1. PBC is a rare disease, with incidence and prevalence reported at variable rates (0.33 to 5.8 100,000/year for incidence; 1.91 to 40.2 per 100,000 for prevalence). Whereas an increase in the incidence has been reported for the last decades, newer global data also indicate changes in the diagnosis and course of the disease (irrespective of treatment) with older ages at diagnosis, and slower progression over time [30]. The pathogenesis of the disease is not fully understood, with environmental, infectious, and genetic predispositions, and with an inflammatory process targeting biliary epithelial cells, and resulting changes of bile-acid metabolism, and enterohepatic circulation being involved.

The disease is characterised with cholestasis, the presence of specific antibodies (AMA and ANA), and histologic evidence of chronic granulomatous, lymphocytic small bile duct cholangitis. The disease course is progressive ultimately leading to the presence of cirrhosis, and end-stage liver disease. In

most patients the progression is slow. In an important subgroup of typically younger patients progression can be more rapid. Usually, PBC is diagnosed on the basis of incident, routine liver transaminase testing at an early stage of disease, without relevant symptoms being present. In addition to the symptoms associated with end-stage liver disease (where present) patients can experience significant systemic symptoms throughout the disease course. Fatigue and pruritus are the most prominent of these symptoms and can be debilitating.

PBC is diagnosed in clinical practice based on the findings of careful history taking, exclusion of other immune-mediated diseases, and the presence of specific findings in imaging, and finally, serological tests, including ANA and AMA. Liver biopsy with histology – according to the current European Practice Guideline – is only recommended in cases with ongoing unexplained cholestasis[31].

PBC is the only disease in this reflection paper, for which a standard therapy is available. Ursodeoxycholic acid (UDCA) is currently the established standard of treatment and has been introduced in the 1990s. More recently, obeticholic acid has been licensed in 2016 for the "treatment of primary biliary cholangitis (previously also known as primary biliary cirrhosis) in combination with ursodeoxycholic acid (UDCA) in adults with an inadequate response to UDCA or as monotherapy in adults unable to tolerate UDCA".

4.3.2. Selection of patient populations

The similar general principles of the selection of study population are also applicable to PBC (see 4.2.2). The specific requirements concern the following:

Because a standard first-line therapy option, plus a second-line add-on therapy option are available, the inclusion of an adequate patient population depends on the intended place in therapy of the investigational agent.

Trials trying to establish a new first-line compound in the disease can include, both, newly diagnosed and/or untreated patients, as well as patients already treated with UDCA. In the latter case, these patients could either have normal liver chemistry (=full responders) (including ALP and bilirubin), or, more adequately, have documented partial response to the agent, which is, however, below the threshold for "unsatisfactory response in one of the established criteria for UDCA response in the literature. The choice of the response criteria will have to be justified. Treatment naive patients should be included based on an unequivocal diagnosis of PBC associated with an (at least) minimal increase in the serological markers of the disease, especially ALP with or without (conjugated) bilirubin elevation, allowing for relevant improvements.

Whereas for early-stage trials, the omission of a histological evaluation, including screening as well as endpoint evaluation is considered acceptable, the availability of a baseline histology evaluation (as well as follow-up evaluation, see below), is highly recommended.

Trials aiming to justify the add-on treatment on top of UDCA will need to include patients based on an insufficient response to UDCA (for the present time, wherein UDCA is the prevailing therapy). A variety of options has been proposed to define such a population, including the so-called Barcelona, Paris-I, -Toronto, Rotterdam, as well as Paris-II criteria [32333435]. However, all these criteria were set-up in order to define a population having the best prognosis at long-term follow-up, and not in order to determine, which of these might delineate a population at the highest risk of progression, and thus be most suitable for additional therapy. An analysis of these different proposals, however, has shown that the likelihood to develop endpoints (such as cirrhosis, decompensation events, and liver transplantation and death, see below) during the course of a trial largely depends on the strictness of these inclusion criteria [36]. It is therefore recommended that the more strict criteria are chosen, allowing only those

patients into the trial which have still a relevant alteration of the serological markers of PBC. Currently, best appears to be the combined use of the ALP≥2xULN, and bilirubin >1xULN despite an at least 1 year therapy with UDCA at the standard recommended dose (10-15 mg/kg b.w./day). Additional criteria with regard to transaminases, albumin, GGT, or Mayo risk score may be applied, if adequately justified.

Trials in the add-on-setting may also include patients not tolerating the standard treatment with UDCA. However, it is expected that these form a minority of patients only in these trials. Nevertheless, consistency of the results needs to be demonstrated.

4.3.3. Study design and endpoints

First-line therapy as alternative to UDCA

Ursodeoxycholic acid (UDCA) has an established efficacy and an acceptable safety record in the disease to be treated. Therefore, the development of alternatives to the first-line therapy have to take into account the level of efficacy, as well as of safety of the standard therapy.

The conduct of such trials may include in addition to a direct comparison to UDCA, also a (potentially small; e.g. based on unequal randomisation) placebo arm for assay sensitivity purposes in case non-inferiority will be the aim of such trials. While it is acceptable to demonstrate non-inferiority to the established treatment, as well as an acceptable safety profile for licensing, considering the properties of the current standard of care with moderate efficacy and relatively good and established safety profiles, it might be necessary to aim at superiority in such trials in order to allow a more clear positive conclusion on risk-benefit, especially in case the safety profile does not allow a conclusion on a similar level of acceptability as for UDCA.

For an intermediate endpoint evaluation strategy in the first line treatment of PBC, as an alternative to UDCA, one of the previously mentioned response criteria (Barcelona, Paris, Rotterdam etc. criteria; see Chapter 4.3.2) can be used, depending on the included population.

In case the untreated population is mainly or solely used, the most obvious endpoint would be the composite of the normalisation of the relevant serological markers, mainly ALP and (total) bilirubin (as composite at the individual level). Any deviation from this stringent definition should be justified.

The trial duration would need to be at least 1 year, with extended (controlled) follow-up (see below) to be planned.

An evaluation of all potential long-term outcomes is considered to be hardly possible in this population, which would be expected to have a high rate of normalisation of the serological markers at the end of the (primary) observation period, and thus have an even delayed further development of disease deterioration. The necessary follow-up treatment documentation would therefore need to demonstrate a prolonged superiority (or at least non-inferiority) for at least 2 years (potentially to be submitted post-licensing) in the serological (Interim) endpoints, supported by an adequate battery of secondary evaluations, based on non-invasive imaging, additional biomarkers, as well as histology. The clinical relevance of these endpoints should be substantiated.

Add-on therapy to UDCA

The reduction of total bilirubin, as well as for ALP (including % reductions and reductions under certain thresholds) have previously been used and accepted as primary endpoints in trials in the add-on setting. These endpoints have to be regarded to be acceptable intermediate outcomes of efficacy in PBC, because currently, it has only been demonstrated for the natural history as well as for UDCA, that the reduction of these two serological markers leads to an overall improved outcome with regard to the

development of end-stage liver disease, decompensation, liver transplantation and death [37]. Hence, an endpoint based on these serological markers is considered an adequate intermediate strategy, which, however, for new compounds, would need to be supported by additional long-term outcomes. The choice of adequate thresholds for the definition of response would need to be adapted to the chosen inclusion criteria, but usually, the most clear-cut thresholds close to normalisation would be expected to be evaluated. Previously, the criteria of ALP<1.67xULN, ALP decrease of at least 15%, as well as (total) bilirubin ≤ULN have been thought to be acceptable. However, more stringent definitions of response are advocated here, with the ALP criterion being at least ALP<1.5xULN with an at least 40% decrease, and total bilirubin ≤ULN. Additional criteria with regard to transaminases, GGT, and/or Mayo score may be added, depending on the respective inclusion criteria [38].

Because the validity of these intermediate endpoints is not fully established, it would usually be expected that long-term outcome data with respect to the histological manifestation of cirrhosis, the decompensation of cirrhosis, MELD score above a threshold defining a high risk of liver related death (e.g. above 14), as well as liver transplantation and death should form the basis for a long-term follow-up evaluation of efficacy. However, the availability of obeticholic acid as first add-on therapy in PBC on the market makes the conduct of placebo-controlled trials with these long-term outcomes more complicated, and adequate escape procedures may be necessary to be implemented into, in order to allow the ethically acceptable conduct of, such trials in the future.

Due to the fact that it is currently not known whether such trials on long-term endpoints finally turn out to be feasible in the disease, the fact that the disease is rare, and the development of later stage disease is slow, the applicants will also have to take care that the best possible evidence with regard to secondary evaluations is also available at the point of interim data evaluation (for the serological endpoint). This should include, but is not restricted to, non-invasive measurements of liver fibrosis/stiffness, serological markers of inflammation and liver damage, as well as histology, including the staging and grading of the disease. The latter item would need to be handled with caution due to the fact that a fully validated histological scoring system for the disease is not available. Historically, Ludwig and Scheuer's classifications, as well a METAVIR have been used in this context, and specific scoring systems seem to be under development[39]. An early consultation within a Scientific Advice procedure is therefore recommended.

Trial durations from 1-2 years have previously been proposed in order to show efficacy on the interim endpoint proposed. From an overall efficacy and safety point of view, but depending on the magnitude of effect to be expected, a study duration of at least 2 years seems to be desirable. A trial extension for the longest possible extend should be aimed at. If indeed studies using long-term outcomes (liver transplantation and death, decompensation events) are intended, these are usually event driven, and a priory determination of the trial duration will not be possible.

It has been proposed that – due to potential ethical concerns with regard to prolonged placebo treatment, as well as adherence problems – that an open-label extension should be conducted (e.g. additional 2 years), and the results could be compared to an external natural history cohort derived from the Global PBC Study Group" database"[40]. However, this is currently not recommended as acceptable strategy and must – for the time being – be also considered as supportive endpoint only.

Despite the availability of at least one alternative add-on treatment at this point of time, the trials in the add-on setting are recommended to be conducted with placebo-control only. This is related to the safety profile of the potential alternative obeticholic acid, which is potentially leading to relevant un-blinding (high occurrence rate of pruritus), and the currently unconfirmed (in respect to long-term outcomes) efficacy status of the compound.

Target of estimation (estimand)

According to ICH E9(R1) Draft Addendum on estimands and Sensitivity Analysis in Clinical Trials (EMA/CHMP/ICH/436221/2017), the scientific question of interest should be specified, and trial features should be aligned.

Potential intercurrent events to be taken into account for the outcome in the setting of PBC can be assumed to be lack of adherence to treatment and the intake of rescue medication. The intake of rescue medication should be considered as a treatment failure (expected to occur in first-line settings), i.e. the composite strategy as discussed in the addendum is considered appropriate for this intercurrent event. With regard to other intercurrent events, a treatment policy strategy appears most suitable, i.e. the outcome regardless of the intercurrent event is of primary interest. Therefore, data on outcome should be collected independently from the occurrence of these intercurrent events, which is considered to be feasible especially in this setting because the primary endpoint(s) is/are based on simple blood biomarker evaluations.

With regard to the evaluation of long-term endpoints it is referred to the respective paragraph on NASH (see 4.2.3).

4.4. Primary sclerosing cholangitis

4.4.1. Short characterisation of the disease

Primary sclerosing cholangitis (PSC) is a rare, chronic, heterogeneous, and idiopathic inflammatory disease characterised by intra- and/or extrahepatic stricturing of bile ducts and development of fibrosis. The natural history of PSC includes the development of complications (e.g. bacterial cholangitis), progression of fibrosis to cirrhosis, and ultimately end-stage liver disease with decompensation, liver transplantation, or death. The disease is frequently associated with inflammatory bowel disease (IBD), including both, Crohn's disease (CD), as well as ulcerative colitis (UC). Patients with PSC are at high risk of cholangiocarcinoma and gall bladder cancer, and also have increased risks of colon carcinoma, whereas the presence of an increased risk for hepatocellular carcinoma is controversial [41][42][43].

Patients are either diagnosed on the presence of cholestasis when screening at risk patients (e.g. those with IBD), or general health screening. Symptoms usually develop with progression of the disease, and include fatigue, pruritus, and right upper quadrant pain, potentially accompanied by jaundice in later stages. The diagnosis is made based on the serum markers of potential cholestasis and finally on the presence of stricturing cholangiopathy, usually diagnosed with magnetic resonance cholangiopancreatography (MRCP). A final diagnosis also requires the exclusion of relevant secondary cholangitis, particularly IgG4– related disease. Liver biopsy is not regularly performed but regarded to be needed in patients with suspected small duct PSC or in patients with suspected overlap with autoimmune hepatitis.[44][45]. The age at diagnosis is mostly between 30 and 40 years, but even children can be affected.

There are a number of factors relevant for the overall prognosis in patients with PSC: The presence of small duct PSC, and of Crohn's disease are associated with a better outcome, whereas ulcerative colitis, and the occurrence of a so-called dominant stricture are factors associated with negative outcomes (with regard to transplant-free survival)[46].

The incidence of the disease has been estimated up to 0.4 to 2.0 per 100,000 inhabitants per year with a wide variability, even within Europe. The prevalence has been estimated to be overall less than 50 per 100,000 (10 per 100,000 inhabitants)[47]. The development of the disease is slow, and it has most

559 recently been estimated that the development of end-stage liver disease may regularly take more than
560 20 years[48].

561 4.4.2. Selection of patient populations

562 Similar principles of the selection of study population as above (see 4.2.2 and 4.3.2) are also
563 applicable to PSC. The specific requirements concern the following:

564 As mentioned above, the diagnosis of PSC mainly relies on the profile of elevated ALP and an abnormal
565 cholangiography consistent with PSC, as shown by MRCP, endoscopic retrograde
566 cholangiopancreatography (ERCP), or percutaneous transhepatic cholangiography (PTC). A minimal
567 threshold for ALP and transaminases elevations should be defined. The availability of a liver biopsy
568 consistent with PSC is a compulsory requirement. The presence of overlap (e.g. with AIH) syndromes
569 can be allowed in exploratory clinical trials, but not in confirmatory trials. Other secondary reasons for
570 PSC also need to be excluded.

571 The selection of PSC patients should allow the occurrence of relevant events in the population included.
572 The inclusion of small duct PSC patients would therefore usually require the presence of other,
573 negative risk factors, such as ulcerative colitis etc. If patients have already a dominant stricture at the
574 time of inclusion, patients should not have relevant fluctuations of serum markers historically, and not
575 have relevant cholestasis at inclusion. It may also be sensible to define an upper limit of other markers
576 of liver damage (e.g. for transaminases). The presence of cholangio- and gall-bladder carcinoma, as
577 well as colon carcinoma should be excluded. Usually, patients with active IBD should not be included
578 due to the potential interference with the search for effective medication, and its associated changes.
579 Similarly, concomitant medication for IBD would require to be stable for a relevant time-frame.
580 Occurrence of acute cholangitis should not have occurred for a relevant time-frame, and no concurrent
581 antibiotic therapy should be part of the medication at inclusion.

582 Depending on the aim of the trial (see Chapter 4.4.3. and 4.5), the inclusion of patients having a
583 relevant level of symptoms should be considered. For a disease modifying study, both, symptomatic,
584 as well as asymptomatic patients, can be included. Despite the fact that UDCA is not regarded as
585 recommended medication in PSC, it is in widespread use. Therefore, the inclusion of patients on
586 concomitant UDCA can be allowed, but intake of UDCA should not be altered during the trial.

587 It would usually be expected that the presence of decompensation symptoms should be an exclusion
588 criterion, but cirrhotic patients without signs and/or symptoms of decompensation can be included.

589 4.4.3. Study design and endpoints

590 No licensed treatment in PSC is currently available. Therefore, a development strategy aiming at the
591 demonstration of effects at an early time-point using intermediate endpoints, for which the surrogacy
592 has at least been demonstrated by natural history studies, with a later confirmation on long-term
593 endpoints, is regarded to be an acceptable option (See also Chapter 4.1.). Such a strategy is also
594 supported by the fact that previous trials in PSC with UDCA have not demonstrated clear beneficial
595 effects for the long-term endpoints, despite being partly successful with potential surrogates.

596 These trials with UDCA in PSC have been assumed to be largely underpowered, and were – although
597 having demonstrated dose-dependent reductions in ALP –not able to demonstrate relevant effects on
598 the long-term efficacy outcomes such as manifestation of cirrhosis, decompensation clinical events,
599 liver transplantation and death. Whereas, however, the level of ALP – at diagnosis and after follow-up
600 – has repeatedly been demonstrated to be associated with outcomes in PSC[49][50], there was an obvious
601 dissociation of ALP and relevant clinical outcomes in the UDCA trials[51][52]. The International PSC Study

Group has recently made comprehensive suggestions for the use of intermediate endpoints in PSC[53], which has been reflected and partially adopted in the regulatory environment [54].

Therefore ALP can currently not be accepted as the only intermediate endpoint to be used in this disease. Other endpoints proposed (such as transient elastography and bilirubin) face similar problems as ALP, or have a less robust history of validation. The use of histology in PSC has been discussed controversially (see also 4.1.), however newer research has been shown that – in addition to its obvious face validity – histology can well be used to evaluate the changes.

Therefore, a combined use of histology evaluation and ALP changes are regarded to represent an acceptable intermediate endpoint for the disease for the time being.

It is again emphasized that intermediate endpoints used for marketing authorisation must be sufficiently reliable to allow the conclusion of a positive benefit risk at time of marketing authorisation. Therefore a co-primary evaluation of these intermediate endpoints should be aimed at. Furthermore it is suggested that a responder-type evaluation based on the criteria of therapeutic response should be the basis, defining serological response as a reduction of ALP to 1.3xULN, or a combination of the reduction to 1.5-1.3xULN with an at least 40% reduction from baseline. For the histological evaluation – best to be based on the newer staging system according to Nakanuma [55] – a similar responder-type evaluation is proposed. The response should be defined based on an at least 1 point improvement in the fibrosis stage. Stable disease (no worsening of fibrosis) could be used instead, if adequately justified.

As advocated before, a later evaluation of long-term outcomes is also considered necessary for PSC , which should be done as a composite endpoint including the manifestation (histological diagnosis) of cirrhosis, a MELD score above 14, decompensation events (such as encephalopathy, variceal bleeding, and ascites), as well as liver transplantation and death.

Due to the slow development of fibrotic stages and the low prevalence of the disease, the difficulties for the validation of the proposed intermediate endpoints are acknowledged. Future applicants should therefore also take care that a sufficient amount of supportive evidence for long-term efficacy is available. This should consist of standard evaluations such as imaging modalities, other biomarkers (bilirubin, transaminases, but also e.g. ELF-test and other potential future biomarkers) as well as important clinical events in the course of the disease, such as (number of) bouts of acute cholangitis, occurrence (manifestation) of dominant stenosis, and finally the occurrence of cholangiocarcinoma, and other malignancies. In case the intended long-term outcome endpoints fail to demonstrate a significant difference to placebo, a final conclusion on the benefit-risk ratio would have to be based on the totality of these data.

As no effective treatment is currently available, the acceptable comparator is regarded to be placebo. Trial duration is anticipated to be 2 years for the interim endpoints, and should be up to 5 years for the demonstration of the long-term clinical outcomes. This proposed trial duration may need modification based on the mechanism of action, as well as anticipated magnitude of effects of new drug candidates, and the fact that usually, an event driven evaluation will be planned for.

Target of estimation (estimand)

Similar to PBC, with reference to ICH E9(R1) the scientific question of interest should be specified, and trial features should be aligned accordingly.

Potential intercurrent events to be taken into account for the outcome in the setting of PSC can be assumed to be lack of adherence to treatment and the occurrence of malignancy. The intake of rescue medication will not play a relevant role for the time being, because no well-established treatments are available.

Also similar to PBC, and according to the character of the primary endpoint, treatment policy strategy may thus be most appropriate for the intercurrent events, i.e. the outcome regardless of the intercurrent event is of primary interest.

4.5. Trials for the symptomatic treatment (PBC and PSC):

It has been described that both, PBC, as well as PSC impose a significant and clinically relevant burden of symptoms on patients with the diseases.

For these reasons, it is possible to develop new treatments in the two diseases, which address the symptomatic improvement of the patients, without aiming generally at positively influencing the natural disease course (disease modification).

Potential drug candidates could involve patients suffering from a variety of symptoms, but at least of the two major features of the disease (fatigue and pruritus). However, if only one symptom of the disease is aimed at, it will usually be expected that effects are also evaluated in other pruritic diseases, in order to claim a general pruritus, or fatigue indication.

If the totality of the disease specific symptoms are aimed at with a treatment, it is recommended that disease specific measurements of the symptoms are part of the primary evaluation. The development of such tools (patient-reported outcome tools – PROs) is encouraged. Usually, a claim of efficacy should be based on an instrument measuring the direct symptoms, supported by a more indirect evaluation of the impact of the symptoms, usually to be evaluated with disease-specific Quality of Life scale.

Clinical trials with this restricted scope could be planned with a limited duration of (placebo) controlled treatment for 6 months. A sufficient amount of long-term data, in order to demonstrate adequate safety should, however, also be available (reference is made to the ICH E1 guideline).

Target of estimation (estimand)

The evaluation of a symptomatic treatment is expected to be evaluated with a treatment policy evaluation. This is partly due to the different character of the endpoints, but also to the partly different nature of the expected intercurrent events, which at least in the case of pruritus could include a variety of rescue treatments. Also, a complete follow-up of patients, even in the case of study drug discontinuation appears to be possible to a higher extent, also supported by the limited observation period.

4.6. Safety considerations

General safety requirements will apply to trials in chronic liver diseases, similar to other fields of drug development. The general requirements to focus on the known pharmacodynamic effects, including off-target effects known from early development programme will fully apply. The following paragraphs therefore deal with the specifics of safety evaluation with regard to liver in patients with underlying liver disease, and the cardiovascular safety consideration applicable to NASH

4.6.1. Safety in PBC and PSC

The underlying liver disease, as well as fluctuations and flares occurring during the course of clinical trials may hamper the evaluation of hepatic safety due to the overlap in accompanying symptoms, as well as the changes in the routine liver safety biomarkers used, such as transaminases, ALP, and bilirubin. The distinction of fluctuation and flare of the underlying disease, from subclinical liver damage and true drug-induced liver injury (DILI) caused by an investigational agent is therefore the

most important feature of the evaluation of liver safety in both disease entities. The distinction of the type of injury pattern, as well as causality assessment (e.g. using the well-established Roussel Uclaf Causality Assessment Method (RUCAM) criteria, as well as expert adjudication), and the search for and potential identification of Hy's law cases are valuable parts of the evaluation of liver safety and potential DILI in clinical trials. In addition, obtaining biopsies whenever possible should be the aim. [5657].

Although a generally increased risk of DILI in patients with underlying liver disease appears to be controversial [58] and may depend on the underlying disease[59], in addition to these general requirements a need exists to define different rules for the safety evaluation during, and after clinical trials with underlying liver diseases. These alternative approaches may include stopping rules, as well as thresholds to define clinically relevant events and the use of novel statistical approaches specifically developed for this purpose[60]. In addition, the inclusion of experimental biomarkers is highly recommended for trials in patients with underlying liver disease [61], but the influence of the underlying disease on these markers should be known before they are used to help the assessment of safety. It is recommended that all these methods are implemented in addition to the routine liver safety evaluation.

4.6.2. Safety in NASH

Similar to PBC and PSC, the evaluation of liver safety in the field is considered paramount, and at the same time, hampered by the underlying disease process. The principles outlined for PBC and PSC are therefore also applicable in NASH. The higher number of patients that can be expected to be treated might, however, allow more clear conclusions on liver safety.

Because NASH is associated with the obesity epidemic, and the liver manifestation of the so-called metabolic syndrome, the patient population included in clinical trials in NASH will be prone to increased risks of adverse events related to concomitant diseases such as arterial hypertension, diabetes mellitus, severe obesity, and hypercholesterolaemia with the associated sequelae cardiovascular events, such as myocardial infarction, stroke, and associated death [626364].

Therefore, depending on the mechanism of action, and the pre-clinical data showing potential detrimental effects with regard to cardiovascular safety, the principles of the "reflection paper on assessment of cardiovascular safety profile of medicinal products" (EMA/CHMP/505049/2015), are considered applicable to NASH also, although it is currently not fully clear whether the risk increase for cardiovascular outcomes and the resulting number of events will allow reliable conclusions. Further long-term natural history data, and long-term clinical trials in the field are needed to draw a final conclusion.

It is therefore necessary, not only to focus the safety evaluation on the occurrence of the so-called major cardiovascular events (MACE) but also on the off-target effects of the potential investigational products on parameters potentially influencing the overall cardiovascular risk, such as plasma lipids, glucose homeostasis, and (systemic) inflammatory parameters.

4.7. Children and adolescents

4.7.1. NASH in children and adolescents

Similar to other aspects of the obesity/"metabolic syndrome" epidemic, non-alcoholic fatty liver disease (NAFLD), as well as NASH have been identified to present an increasingly significant health burden in children and adolescents. The prevalence of NAFLD in children is estimated to be around 10-14%

depending on age. Whereas 2-4 year old children are expected to suffer from NAFLD at only very low rates, the prevalence in adolescents almost reaches adult levels [65].

Assuming a similar rate of patients developing NASH from the presence of NAFLD as in adults [66] , it is clear that NASH is a relevant health problem also in the young age group, although the development of late-stage disease may take years and might be expected to manifest not before reaching adulthood. However, rapid progression to advanced liver disease in childhood has been described[67].

Therefore, there is a relevant medical need to develop treatments for NASH also in children.

As outlined above, the diagnosis of NASH is currently considered to require the conduct of liver biopsy with histological evaluation, and the conduct of clinical trials should be mainly based on repeated biopsy results. The diagnosis itself is also based on histology in childhood/adolescence patients[6869].However, the conduct of repeated biopsies in clinical trials requires increased awareness of the potentially associated ethical and procedural problems when children are concerned, and the need for non-invasive outcomes in this population is therefore considered to be of even higher priority.

Furthermore, the histology evaluations available have shown distinct features of paediatric NASH as compared to adults, with the presence of a relevant proportion of patients developing a unique histology with presence of portal-based chronic inflammation (and fibrosis) (as opposed to the lobular inflammation found in adults and less ballooning[70]). The clinical meaningfulness of this distinct type of histology in children is currently unknown, and consequently, a different histological scoring system may be needed for the paediatric population.

The development of new medicinal products for the treatment of NASH in children therefore requires first of all the collection of new and evaluation of existing data with regard to the natural history of the disease.

Drug development in children will also require the final determination of the adequate age range to be studied. Young children (e.g. below the age of 6-10 years), might still be early in the disease process, and therefore be appropriate candidates for non-pharmacological interventions, such as life-style and dietary changes, of which success rates (with regard to weight loss) are usually higher than in adults. Consequently, the potential for regression of inflammatory changes is similarly considered to be higher[71].

The development of new medicinal products for NASH in children would also need a determination of the quantity of data needed to be available for adults, before therapeutic trials are conducted. At this point of time – when there still seems to be a need for more natural history data – it is recommended that relevant clinical trials are deferred until data in adults on long-termer endpoints are available (with regard to progression to cirrhosis, liver transplantation and death) at least until the validity of the proposed interim endpoints has been relevantly substantiated.

The availability of further data on natural history, as well as on the individual new compound in adults might already enable to more precisely determine the level of extrapolation that can be applied (see draft: Reflection paper on the use of extrapolation in the development of medicines for paediatrics. EMA/199678/2016).

Once the above mentioned data are available, and a decision on the possible level of extrapolation can be taken, the conduct of therapeutic trials in children is considered to be relevant, keeping in mind the potential for enhanced regression of NASH. Besides the necessary investigation of the appropriate dose (under full consideration of the potential differences in pharmacokinetics in obese and NASH adolescents compared to adults), and development of age-appropriate formulations, the conduct of placebo-controlled trials, including endpoints based on histology, and thus, repeated liver biopsies may still be required in order to fully account for the differences between childhood/adolescent and adult

NASH. Even if from adult studies, an intermediate endpoint method such as an early histology evaluation endpoint, imaging methods, or biomarkers, have partly been validated, it can be anticipated that these would have to undergo further validation in children

The conduct of studies with histology endpoints should take full account of the potential for the ethical problems associated with any more than minimally invasive procedures, and may need a careful approach with regard to the patient selection (e.g. older age groups, more advanced disease, etc.).

4.7.2. PBC: Children and adolescents

The youngest reported age of a confirmed disease onset has been in a 15-year old post-menarche adolescent [72] , and it is thought that a true paediatric disease is not encountered.

Potential applicants developing new substances in the treatment of PBC would therefore be expected to apply for a waiver for a paediatric programme in the disease.

4.7.3. PSC in Children and Adolescents

Paediatric PSC is a very rare disease, even compared to adult PSC, which itself is classified as orphan. However, it is estimated that the risk in patients with IBD to develop PSC is doubled in the paediatric population as compared to adults. Therefore, PSC appears to be a major source of morbidity in this population. With the rising incidence of IBD, a clear unmet medical need exists. Also distinct from adult PSC, there is a higher overlap of PSC with other syndromes, especially AIH (PSC-AIH-overlap syndrome or Autoimmune Sclerosing Cholangitis - ASC) [73][74] . The investigation of new compounds, also for children is therefore considered to be needed.

Although a relevant amount of data has already been collected for paediatric PSC [75], there is still a need to collect further natural history data before clinical trials in PSC can reasonably be undertaken.

Once these natural history data are available and have been evaluated to a sufficient extend, trials in paediatric PSC may also be conducted with patients suffering from overlap conditions (especially AIH-PSC), if adequate. The inclusion of patients should be based on the identified risk factors, which are distinct from adult PSC, such as elevated gamma-glutamyl-transferase (GGT) and aspartate aminotransferase-to-platelet ratio index (at diagnosis).

Besides the need to fully explore the PK profile in the respective population, there can currently no clear recommendations be given with regard to the design of trials, and endpoints to be used. Consultation with the agency early in the drug development (scientific advice and submission of PIP) is therefore advisable.

5. References

[1] Sanyal AJ et al: Challenges and opportunities in drug and biomarker development for non-alcoholic steatohepatitis: Findings and Recommendations from an American Association fo the Study of Liver Disease – US Food and Drug Administration joint workshop. Hepatology 2015; 61: 1392-1405.
[2] Ratziu V e tla: Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology 2005; 128: 1898-1906.
[3] Seeff LB et al: Complication Rate of Percutaneous Liver Biopsies among Persons with Advanced Chronic Liver Disease in the HALT-C Trial. Clin Gastroenterol Hepatol 2010; 8: 877-883.
[4] Rinella ME et al: Nonalcoholic fatty liver disease. A systematic review. JAMA 2015; 313: 2263-2273.
[5] Younossi ZM et al: Global epidemiology of nonalcoholic fatty liver disease – meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology 2016; 64: 73-84.
[6] Younossi Z et al: Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. Nat Rev Gastroenterol Heptaol 2018; 15: 11-20.
[7] Cholankeril G et al: Liver Transplantation for non-alcoholic steatohepatitis in the US: Temporal trends and outcomes. Dig Dis Sci 2017; 62: 2915-2922.
[8] Eslam M et al: Genetics and epigenetics of NAFLD and NASH: Clinical impact. J Hepatol 2018; 68: 268-279.

[9] Bellentani S: The epidemiology of non-alcoholic fatty liver disease. Liver International 2017; 37 (Suppl. 1) 81-84.
[10] European Association for the Study of the Liver (EASL), European Association for the Study of Diabetes (EASD), European Association for the Study of Obesity (EASO): EASL-EASD-EASO Clinical practice guidelines for the management of non-alcoholic fatty liver disease. Diabetologia (2016) 59: 1121-1140.
[11] Dan AA et al: Health-related quality of life in patients with non-alcoholic fatty liver disease. Aliment Pharmacol Ther 2007; 26: 815-820.
[12] Sattar N et al: Non-alcoholic fatty liver disease. BMJ 2014; 349: g4596 doi: 10.1136/bmj.g4596
[13] Dulai P et al: Increased risk of mortality by fibrosis stage in non-alcoholic fatty liver disease: Systematic review and meta-analysis. Hepatology 2017: 65t: 1557-1565.
[14] Marengo A et al: Progression and natural history of non-alcoholic fatty liver disease in adults. Clin Liver Dis 2016; 20: 313-324
[15] Goh GBB and AJ McCullough: Natural history of non-alcoholic fatty liver disease. Dig Dis Sci 2016; 61: 1226-1233.
[16] Singh S et al: Fibrosis progression in non-alcoholic fatty liver vs non-alcoholic steatohepatitis: A systematic review and meta-analysis of paired-biopsy studies. Clinical Gastroenterology and Hepatology 2015; 13: 643-654.
[17] Hagström H et al: Fibrosis stage but not NASH predicts mortality and time to development of severe liver disease in biopsy-proven NAFLD. J Hepatol 2017; 67: 1265-1273.
[18] Kleiner DE et al.: Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;41:1313–21.
[19] Siddiqui MS et al: Case definition for inclusion and analysis of endpoints in clinical trials for non-alcoholic steatohepatitis through the lens of regulatory science. Hepatology 2018; 67: 2001-2012.
[20] Patel Y et al: Baseline Parameters in clinical trials for non-alcoholic steatohepatitis: Recommendations from the Liver Forum. Gastroenterology 2017; 153: 621-625.
[21] Boursier J et al: Screening for therapeutic trials and treatment indication in clinical practice: MACK-3, a new blood test for the diagnosis of fibrotic NASH. Aliment Pharmacol Ther 2018; 47: 1387-1396.
[22] Vilar-Gomez E and N Chalasani: Non-invasive assessment of non-alcoholic fatty liver disease: Clinical prediction rules and blood-based biomarkers. J Hepatol 2018; 68: 305-315.
[23] Vilar-Gomez E et al: Weight Loss through lifestyle modification significantly reduces features of nonalcoholic steatohepatitis. Gastroenterology 2015; 149, 367-378.
[24] Ratziu V: A critical review of endpoints of non-cirrhotic NASH therapeutic trials. J Hepatol 2018; 68: 353-361.
[25] D'Amico G et al: Natural history and prognostic indicators of survival in cirrhosis: A systematic review of 118 studies. J Hepatology 2006; 44: 217-231.
[26] Ripoll C et al: Hepatic venous pressure gradient predicts clinical decompensation in patients with compensated cirrhosis. Gastroenterology 2007; 133: 481-488.
[27] Polyzos SA et al: Nonalcoholic fatty liver disease: Is it time for combination treatment and a diabetes-like approach. Hepatology 2018; 68: 389-389
[28] Beuers U et al: Changing nomenclature for PBC: from "cirrhosis" to "cholangitis". Gut 2015; 64: 1671-1672.
[29] Carey EJ et al: Primary biliary cirrhosis. The Lancet 2015; 386; 1565-1575.
[30] Murillo Perez, CF et al: Milder disease stage in patients with primary biliary cholangitis over a 44-year period: A changing natural history. Hepatology 2018; 67: 1920-1929.
[31] European Association for the Study of the Liver: EASL Clinical Practice Guidelines: The diagnosis and management of patients with primary biliary cholangitis. Hepatology 2017; 67: 145-172.
[32] Pares A et al: Excellent long-term survival in patients with primary biliary cirrhosis and biochemical response to ursodeoxycholic acid Gastroenterol 2006; 130: 715-720.
[33] Kumagi T et al: Baseline ductopenia and treatment response predict long-term histological progression in primary biliary cirrhosis. Am J Gastreoenterol 2010; 105: 2186-2194.
[34] Kuiper EMM et al: Improved prognosis of patients with primary biliary cirrhosis that have a biochemical response to ursodeoxycholic acid. Gastroenterology 2019; 136: 1281-1287.
[35] Corpechot C et al: Early primary biliary cirrhosis: biochemical response to treatment and prediction of long-term outcome J Hepatol 2011; 55: 1361-1367.
[36] Momah N et al: Optimizing biochemical markers as endpoints for clinical trials in primary biliary cirrhosis. Liver Int 2012; 32: 790-795
[37] Lammers WJ et al: Levels of alkaline phosphatase and bilirubin are surrogate end points of outcomes of patients with primary biliary cirrhosis: An international follow-up study. Gastroenterol 2014; 147: 1338-1349.
[38] Corpechot C et al: A placebo-controlled trial of bezafibrate in primary biliary cholangitis. NEJM 2018; 378:2171-2181.
[39] Wendum D et al: Primary biliary cirrhosis: proposal for a new simple histological scoring system. Liver International 2015; 35: 652-659.
[40] Lammers WJ et al: Development and validation of a scoring system to predict outcomes of patients with primary biliary cirrhosis receiving ursodeoxycholic acid therapy Gastroenterology 2015; 149: 1804-1812.
[41] Dyson JK et al: Primary sclerosing cholangitis: The Lancet 2018; 391:2547-2559
[42] Karlsen TH et al: Primary sclerosing cholangitis – a comprehensive review J Hepatol 2017; 67: 1298-1323.
[43] Horley-Silva JL et al: An update on cancer risk and surveillance in primary sclerosing cholangitis. Liver Int 2017; 37: 1103-1109.
[44] European Association for the Study of the Liver. EASL clinical practice guidelines: management of cholestatic liver disease. J Hepatol 2009; 51: 237-267.
[45] Lindor KD et al: ACG Clinical Guideline: Primary Sclerosing Cholangitis. Am J Gastroenterol 2015; 110: 646-659.[46] Weismüller, TJ et al. Patient age, sex, and inflammatory bowel disease phenotype associate with course of primary sclerosing cholangitis. Gastroenterology. 2017; 152: 1975–1984
[47] Molodecky NA et al: Incidence of primary sclerosing cholangitis: a systematic review and meta-analysis. Hepatology 2011; 53: 1590-1599.
[48] Boonstra K et al: Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis; Hepatology2013; 58: 2045-2055.
[49] Rupp, C et al: Reduction in alkaline phosphatase is associated with longer survival in primary sclerosing cholangitis, independent of dominant stenosis. Aliment Pharmacol Ther 2014; 40; 1292-1301

[50] De Vries EMG et al: Alkaline phosphatase at diagnosis of primary sclerosing cholangitis and 1 year later: evaluation of prognostic value. Liver International 2016; 36: 1867-1875.
[51] Olsson R et al: High-Dose Ursodeoxycholic Acid in Primary Sclerosing Cholangitis: A 5-Year Multicenter, Randomized, Controlled Study. Gastroenterology 2009; 129: 1464-1472,
[52] Lindor KD et al: High-dose ursodeoxycholic acid for the treatment of primary sclerosing cholangitis. Hepatology 2009; 50: 808-814.
[53] Ponsioen CY et al: Surrogate endpoints for clinical trials in primary sclerosing cholangitis: Review and results from an International PSC Study Group consensus process. Hepatology 2016; 63; 1357-1367,
[54] Ponsioen CY et al: Design and Endpoints for Clinical Trials in Primary Sclerosing Cholangitis. Hepatology 2018. E-pub ahead of print. Hepatology 2018; 68: 1174-1188
[55] De Vries EMG et al: Validation of the prognostic value of histologic scoring systems in primary sclerosing cholangitis: An international cohort study. Hepatology 2017; 65: 907-919
[56] Teschke R and G. Danan: Diagnosis and management of drug-induced liver injury (DILI) in patients with pre-existing liver disease. Drug Saf 2016; 39: 729-744.
[57] Kullak-Ublick GA et al: Drug-induced liver injury: recent advances in diagnosis and risk assessment. Gut 2017; 66: 1154-1164
[58] Teschke R and G Danan: Drug-induced liver injury: Is chronic liver disease a risk factor and a clinical issue? Exp Op Drug Metabol and Toxicol 2017; 13: 425-438.
[59] Massart J et al: Role of nonalcoholic fatty liver disease as risk factor for drug-induced hepatotoxicity. J Clin Transl Res 2017; 3: (Suppl 1) 212-232.
[60] Kullak-Ublick GA et al: Liver safety assessment in special populations (Hepatitis B, C, and oncology trials). Drug Saf 2014 (Suppl 1) 37: S57-62.
[61] Letter of support for drug-induced liver injury (DILI) biomarker. EMA/423870/2016
[62] Adams LA et al: Non-alcoholic fatty liver disease and its relationship with cardiovascular disease and other extrahepatic diseases. Gut 2017; 66: 1138-1153.
[63] Leonardo A et al: Hypertension, diabetes, atherosclerosis and NASH: Cause or consequence? J Hepatol 2018; 68: 335-352.
[64] Sookian S and CJ Pirola: Systematic review with meta-analysis: risk factors for non-alcoholic fatty liver disease suggest a shared altered metabolic and cardiovascular profile between lean and obese patients. Aliment Pharmacol Ther 2017; 46: 85-95.
[65] Doycheva I et al: Nonalcoholic fatty liver disease in adolescents and young adults: The next frontier in the epidemic.Hepatology 2016; 65: 2100-2109.
[66] Schwimmer JB et al: Prevalence of fatty liver in children and adolescents, Pediatrics 2006; 37: 1202-1219.
[67] Kohli R et al: Rapid Progression of NASH in childhood. J Pediatr Gastroenterol Nutr 2010; 50: 453-456.
[68] Vajro P et al: JPGN 2012; 54: 700-713.
[69] Vos M et al: NASPGHAN Clinical Practice Guideline for the diagnosis and treatment of nonalcoholic fatty liver disease in children: Recommendations from the expert committee on NAFLD (ECON) and the North American Society of Pediatric Gastroenterology, Hepatology and Nutrition (NASPGHAN), JPGN 2017; 64: 319-334.
[70] Schwimmer JB et al: Histopathology of pediatric nonalcoholic fatty liver disease. Hepatology 2005; 42: 641-649.
[71] Vos M et al (see 73).
[72] Dahlan Y et al: Pediatric-onset primary biliary cirrhosis. Gastroenterol 2003; 125: 1476-1479.
[73] Cotter JM and CL Mack: Primary Sclerosing Cholangitis: Unique aspects of Disease in children. Clinical Liver Disease 2017; 10: 120-123.
[74] Mieli-Vergani, G et al: Diagnosis and Management of Pediatric Autoimmune Liver Disease: ESPGHAN Hepatology Committee Position Statement. JPGN 2018; 66: 345–360.
[75] Deneau MR et al: The natural history of primary sclerosing cholangitis in 781 children: A multicentre, international collaboration. Hepatology 2017; 66: 518-527.