4 December 2023
EMA/CHMP/299976/2018
Committee for Medicinal Products for Human Use (CHMP)

Reflection paper on regulatory requirements for the development of medicinal products for primary biliary cholangitis (PBC) and primary sclerosing cholangitis (PSC)

Draft agreed by Gastroenterology Drafting Group - within the Reflection paper on regulatory requirements for the development of medicinal products for chronic non-infectious liver diseases (PBC, PSC, NASH).	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
Agreed by Rheumatology and Immunology Working Party	July 2023
CHMP	December 2023

Keywords	Primary Biliary Cholangitis (PBC), Primary Sclerosing Cholangitis (PSC), clinical study design, endpoints, conditional marketing authorisation, unmet medical need, liver biopsy

Official address Domenico Scarlattilaan 6 ● 1083 HS Amsterdam ● The Netherlands
Address for visits and deliveries Refer to www.ema.europa.eu/how-to-find-us
Send us a question Go to www.ema.europa.eu/contact **Telephone** +31 (0)88 781 6000 An agency of the European Union

Table of contents

1. Introduction

Primary biliary cholangitis (PBC) and primary sclerosing cholangitis (PSC) are chronic, non-infectious, cholestatic liver diseases. The number of available treatment options for these diseases is still limited, and the development of new medicinal products is challenging due to the slow disease progression and other characteristics of PBC and PSC.

2. Scope

This reflection paper intends to address the EU regulatory position on the main topics of the clinical development of new medicinal products for the treatment of patients with PBC and PSC. For both disease entities, the regulatory experience with licensing of new medicinal products is limited. Therefore, this paper aims at a preliminary definition of development strategies. Additional regulatory decision making is anticipated and will be helpful to replace this reflection paper with a potential full guidance document in the future.

In this reflection paper, development of treatments that concern disease modifying effects are discussed first. Later, symptomatic treatment of cholestatic pruritus in PBC and PSC is discussed (section 6). Scientific advice is recommended in case a particular topic about PBC and/or PSC is not covered in this reflection paper.

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 and regulations, especially the following:

- Reflection paper on the use of extrapolation in the development of medicines for paediatrics. (EMA/189724/2018)
- Guideline on clinical investigation of medicinal products in the pediatric population (ICH E111(1))
- Points to consider on application with 1. Meta-Analyses; 2. One pivotal study. (CPMP/EWP/2330/99).
- ICH E9(R1) Addendum on estimands and sensitivity analysis in clinical trials (EMA/CHMP/ICH/436221/2017)
- Guideline on the evaluation of pharmacokinetics of medicinal products in patients with impaired hepatic function (CPMP/EWP/2339/02)
- Guideline on the choice of the non-inferiority margin (EMA/CPMP/EWP/2158/99)
- ICH E10 Note for guidance on choice of control group in clinical trials (CPMP/ICH364/96)
- Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products (EMEA/CHMP/EWP/139391/2004)
- Guideline on the scientific application and the practical arrangements necessary to implement Commission Regulation (EC) No 507/2006 on the conditional marketing authorisation for medicinal products for human use falling within the scope of Regulation (EC) No 726/2004 (EMA/CHMP/509951/2006)

4. Primary biliary cholangitis

4.1. Short description of the disease

4.1.1. Epidemiology and aetiology

Primary biliary cholangitis (PBC), previously known as primary biliary cirrhosis[1] is a chronic, slowly progressive autoimmune cholestatic liver disease[2]. The disease is mainly diagnosed in female patients with a female : male ratio of about 10:1. PBC is a rare disease, with an 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 in the last decades, newer global data also indicate changes in the diagnosis and course of the disease (irrespective of treatment) with older age at diagnosis, and slower progression over time[3].

The pathogenesis of the disease is not fully understood. Environmental, infectious, and genetic factors predispose to disease development[4]. An inflammatory process targeting biliary epithelial cells, bile-acid metabolism changes, and the enterohepatic circulation are probably involved in the pathogenesis. Around 2-19% of patients with PBC have features of autoimmune hepatitis (AIH), referred to as having AIH-PBC overlap syndrome[5]. Other concomitant autoimmune diseases, such as autoimmune thyroiditis, are prevalent in patients with PBC.

4.1.2. Clinical presentation and disease course

Approximately 50 percent of patients with PBC are asymptomatic at diagnosis and are detected because of abnormalities in liver biochemical tests obtained for other reasons where no such abnormalities are expected. Among patients with symptoms, fatigue and pruritus are most common and can be debilitating. In newly diagnosed patients, approximately half of the patients complain of fatigue and one-third of the patients suffer from pruritus. The incidence and severity of pruritus increases during the course of the disease (see section 6 on symptomatic treatment of cholestatic pruritus). In addition, patients may have signs and symptoms due to concomitant autoimmune disorders. These symptoms can occur at any stage of the disease and do not necessarily correlate with disease severity.

The disease course is progressive, ultimately leading to cirrhosis, and end-stage liver disease. In most patients the progression is slow. In younger patients, progression can be more rapid.

4.1.3. Laboratory measurements and histology

The disease is characterised by cholestasis, typically presenting biochemically by increased levels of alkaline phosphatase (ALP) and gammaglutamyl transferase (GGT) together with the presence of specific auto-antibodies, i.e. anti-mitochondrial antibodies (AMA) and PBC specific antinuclear antibodies (anti-sp100, anti-gp210). Serum bilirubin (a marker of cholestasis and liver function) and transaminases may be elevated at diagnosis, but often indicate a more advanced disease or AIH-PBC

[1] Beuers U et al: Changing nomenclature for PBC: from "cirrhosis" to "cholangitis". Gut 2015; 64: 1671-1672.
[2] Carey EJ et al: Primary biliary cirrhosis. The Lancet 2015; 386; 1565-1575.
[3] 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.
[4] Mantaka A et al: Primary biliary cirrhosis in a genetically homogenous population: Disease associations and familial occurrence rates. BMC Gastroenterology 2012; 12: 110.
[5] Boberg KM et al: Overlap syndromes: the International Autoimmune Hepatitis Group (IAIHG) position statement on a controversial issue. J Hepatol 2011; 54: 374-85.

overlap syndrome. Increased levels of immunoglobulin (IgM) are typically seen in PBC. Histologically there is evidence of chronic granulomatous, lymphocytic small bile duct cholangitis.

Primary biliary cholangitis is diagnosed based on the findings of the above elevated cholestatic liver tests and serological tests and by exclusion of other causes of chronic cholestasis. In clinical practice, liver biopsy with histology is not needed for diagnosing PBC and is only recommended in cases with ongoing unexplained cholestasis or when co-existent AIH is suspected[6].

4.1.4. Current treatment options

The goals of treatment and management are the prevention of end-stage liver disease and the amelioration of associated symptoms. Ursodeoxycholic acid (UDCA) is the established first-line treatment for PBC, introduced in the 1990s. Current second-line treatments have either conditional marketing authorisation (CMA) or are used off-label.

4.2. Objectives of developments of medicinal products for treatment of PBC

The clinical development plan of anew medicinal product can aim at different indications. The inclusion of a relevant patient population depends on the intended place in therapy of the investigational agent. Further, medicinal products may be developed to modify disease progression or may be developed to achieve symptomatic improvement.

The indication sought strongly informs the estimand that is of regulatory interest and the study design. Currently the following types of approaches are foreseen:

- First-line treatment
- Second-line treatment (monotherapy or add-on)
- Symptomatic treatment of cholestatic pruritus

The application of the estimand framework, the study design, and the study duration for the first- and second- line treatments are discussed below. Applications for symptomatic treatment of cholestatic pruritus is discussed in Section 6.

4.2.1. Study design and comparators

4.2.1.1. First-line therapy as alternative to UDCA

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

Studies in the first-line setting can be designed as either a non-inferiority or a superiority study with UDCA as a comparator. For patients already being treated with UDCA, patients should be randomised to either continue treatment with UDCA or to the new treatment. If the study population includes patients with a satisfactory response to UDCA (complete responders), the protocol has to ensure appropriate monitoring and rescue therapy in the case of a loss of response.

[6] 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.

It is acceptable to demonstrate non-inferiority to the established treatment, for establishing a positive benefit-risk balance. If performed in a low-risk population it might be necessary to aim at superiority in such studies in order to allow a more robust conclusion on the benefit-risk balance, especially in case the safety profile does not allow a conclusion on a similar level of acceptability as for UDCA.

4.2.1.2. Second-line therapy (monotherapy or add-on to UDCA)

There exists a medical need for alternative second-line therapies for PBC. Studies in the add-on setting are recommended to be conducted with placebo-control only.

4.2.2. Selection of patient populations

The inclusion of a relevant patient population depends on the intended place in therapy of the investigational agent. It is recommended to study patients with different stages of the disease, i.e. both early PBC and advanced PBC. Patients with AIH-PBC overlap syndrome should normally be excluded from studies by use of relevant criteria for the diagnosis of this sub-group. The relevant patient population should be stated in the 'Population' attribute of the estimand(s) targeted by the study.

4.2.2.1. First-line therapy as alternative to UDCA

Studies 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. Treatment naive patients should be included based on an unequivocal diagnosis of PBC associated with a defined minimal increase in the biochemical markers of the disease, especially ALP with or without bilirubin elevation, allowing for relevant improvements.

Patients already treated with UDCA could either have normal liver chemistry (complete response) or alternatively, they could have a documented partial response to the agent, which is below the threshold for 'unsatisfactory' response in one of the established criteria for UDCA non-response. The choice of the criteria for complete or partial response at baseline will have to be justified based on literature.

The availability of a baseline histology evaluation as well as follow-up evaluation in at least a sub-group of patients is recommended.

4.2.2.2. Second-line therapy in patients unresponsive or intolerant to UDCA

Studies aiming for an indication of add-on treatment on top of UDCA will need to include patients based on an insufficient response to UDCA. Patients intolerant to UDCA may also be included but considering the overall benign safety profile of UDCA are expected to be a minority of the study population. Nevertheless, consistency of the efficacy results for this "intolerant" subpopulation with the results in patients who experienced an insufficient response to study treatment needs to be demonstrated. A variety of options have been proposed to define a non-responsive population,

including the Barcelona, Paris-I, Toronto, Rotterdam, as well as the Paris-II criteria[7] [8] [9] [10].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 definitions, however, has shown that the likelihood to develop endpoints such as cirrhosis, decompensation events, liver transplantation and death during the course of a study largely depends on the strictness of the inclusion criteria[11]. Selection of patients for second-line therapy based on the combined use of ALP≥1.67xULN, and bilirubin within the normal range of < 2xULN despite an at least 1-year therapy with UDCA at the standard recommended dose (12-16 mg/kg/day) is acceptable. However, based on the above, it is recommended that more strict criteria are chosen, allowing only those patients into the study which have still a relevant alteration of the serological markers of PBC. A "high normal range" for bilirubin or abnormal bilirubin may be used as criteria for enrichment with patients being at higher risk of progression. Additional criteria with regard to transaminases, albumin, GGT, or different risk scores may also be applied, if adequately justified.

Non-invasive tests, e.g. for evaluating liver fibrosis stage are recommended to be used for further enrichment and/or stratification. Patients with different stages of the disease should be studied, either in separate studies or in the same study. Consistency of results for patients with more advanced disease compared to those with stable, less advanced disease would be expected.

Further, patients intolerant or unresponsive to second-line treatments can be included, in this case these treatment(s) would then be replaced by the investigative agent. The availability of a baseline histology evaluation as well as follow-up evaluation in at least a sub-group of patients, is highly recommended in patients non-responsive to UDCA.

4.2.3. Endpoints and summary measure

4.2.3.1. First-line therapy as alternative to UDCA

Primary endpoint

In the first-line setting, a relevant endpoint would be the composite endpoint of the normalisation of the relevant biochemical markers, mainly ALP and total bilirubin (as composite endpoint at the individual level). Any deviation from this stringent definition should be justified.

Summary measure

A responder analysis is recommended, i.e. percentage of responders in study drug vs control group based on a pre-defined definition of a responder. For non-inferiority studies, a margin needs to be defined and justified based on natural history data and clinical study experience with UDCA.

[7] 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.
[8] 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.
[9] 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.
[10] Corpechot C et al: Early primary biliary cirrhosis: biochemical response to treatment and prediction of long-term outcome J Hepatol 2011; 55: 1361-1367.
[11] Momah N et al: Optimizing biochemical markers as endpoints for clinical trials in primary biliary cirrhosis. Liver Int 2012; 32: 790-795.

Supportive endpoints

The biochemical markers would need to be supported by other endpoints to allow for an assessment based on the totality of data. A sufficiently powered and meaningful evaluation of all potential long-term liver related clinical outcomes is considered to be hardly possible in this population with slowly progressive disease, also taking into account the setting of an active controlled study. A prolonged superiority or at least non-inferiority of study treatment for a long-term period in the biochemical endpoints, supported by an adequate battery of secondary evaluations, based on e.g. symptoms, non-invasive imaging for liver stiffness evaluation, additional biomarkers, as well as histology will support the totality of evidence. The clinical relevance of each of these endpoints should be substantiated.

For the evaluation of long-term treatment outcomes, it would be expected to have a high rate of sustained normalisation of the biochemical markers at the end of the (primary) observation period, and thus have a delayed further development of disease deterioration.

Study duration

The study duration would need to be at least 2 years, with extended (controlled) follow-up to be planned.

4.2.3.2. Second-line therapy (monotherapy or add-on to UDCA)

Primary endpoint

The reduction of ALP, as well as of total bilirubin (including % reductions and reductions under certain thresholds) have previously been used and accepted as primary endpoints in studies in the add-on setting and could support a CMA. Confirmation of clinical benefit, however, may be difficult. Thus, an alternative regulatory strategy to consider would be to aim for full approval, using pivotal data in patient populations with different stages of the disease.

At present, it has only been demonstrated for the natural history, as well as for UDCA, that the reduction of ALP and bilirubin leads to an overall improved outcome with regard to the development of end-stage liver disease, decompensation, liver transplantation and death[12]. Whereas on one hand a primary endpoint based on these markers is considered acceptable, on the other hand it needs to be supported by additional secondary clinical endpoints. The choice of adequate thresholds for the definition of response would need to be adapted to the chosen inclusion criteria. However, the most clear-cut thresholds close to normalisation are expected to be evaluated. An endpoint of complete response, i.e. normalisation of ALP and a bilirubin level of < 0.7xULN is highly recommended as a primary or secondary endpoint. Use of stringent definitions such as complete response has been shown to significantly reduce the placebo response[13]. A primary endpoint of response may also be defined as at least ALP<1.5xULN with an at least 40% decrease, or ALP<1.67xULN, ALP decrease of at least 15%, provided that other secondary endpoints support the clinical effects of study treatment, including the more stringent definitions of response. Additional criteria with regard to transaminases, GGT, symptoms, and/or selected risk scores may be added, depending on the respective inclusion criteria.

Summary measure

A responder analysis is recommended, i.e. percentage of responders in study drug vs control group based on a pre-defined definition of a responder.

[12] 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.
[13] Corpechot C et al: A placebo-controlled trial of bezafibrate in primary biliary cholangitis. NEJM 2018; 378:2171-2181.

Supportive endpoints

Supportive endpoints include symptomatic response, and endpoints based on non-invasive tests, and histology as well as longer-term liver related outcomes (e.g. decompensation events, signs of portal hypertension). Different definitions of response are relevant secondary endpoints, in particular a more stringent definitions of response may be needed depending on the definition of response for the primary endpoint.

Because the validity of the biochemical markers is not fully established, it would usually be expected that long-term outcome data with respect to the diagnosis of cirrhosis (based on histology or liver stiffness measurement), decompensation events of cirrhosis, MELD score of > 14 defining a high risk of liver related death, as well as liver transplantation and death should form the basis for a long-term follow-up evaluation of efficacy. Measures to improve the quality of data, which will facilitate any assessment of the totality of data in situations where fully powered studies using clinical outcomes are not feasible include use of relevant secondary endpoints, numerical observations of liver related events and sufficient long-term observation time.

Given that it is currently not known whether such studies on long-term endpoints finally turn out to be feasible, 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. This should include, but is not restricted to, non-invasive measurements of liver fibrosis/stiffness, biochemical markers of inflammation and liver function, as well as histology.

Study duration

From an overall efficacy and safety point of view, but depending on the magnitude of effect to be expected, a study durationis anticipated to be approximately 2 years, and for the demonstration of the long-term clinical outcomes a study duration of **several** years is anticipated.

4.2.4. Intercurrent events and strategies

It would be expected that all intercurrent events are pre-defined in the study protocol and compared between the treatment arms. Among potential intercurrent events to be taken into account for the outcome in the setting of PBC (first line and second-line treatment) is intake of rescue medication and other types of lack of adherence e.g. treatment discontinuation, treatment interruption, dosing deviation and intake of concomitant medication. The intake of rescue medication, or liver-related events and death from any cause should be considered as a treatment failure, i.e. the composite strategy as discussed in the addendum is considered appropriate for these intercurrent events. In general, unless an alternative is justified, for other intercurrent events, including those related to lack of adherence to treatment, the "treatment-policy" strategy should be applied, i.e. the outcome regardless of the intercurrent event is of primary interest.

Carefully defining secondary estimands for important secondary endpoints is highlighted in the setting where support from multiple endpoints is essential.

4.2.5. Study design and statistical analysis

Choices made for study design, data collection and statistical analysis should be aligned to the scientific question of interest that is posed by the study objective. This requires a detailed specification of the estimand (the "target of estimation"), including the specification of strategies to handle each of the relevant events that occur after randomisation and that would affect the interpretation of an outcome variable or preclude its observation (intercurrent events).

For intercurrent events that are intended to be accounted for by the treatment policy strategy, data with regard to the outcomes of interest should be collected independently from the occurrence of the intercurrent event, which is considered to be feasible especially in this setting because the primary endpoint(s) is/are based on simple blood biomarker evaluations. Data that is nevertheless not collected, for example in case the patient discontinues the study, results in a missing data problem with regard to subsequent statistical inference. Generally, sensitivity analyses to support the robustness of the primary analysis should be provided.

Considering a patient with missing data as a non-responder usually results in a conservative estimate of the treatment effect for the recommended primary estimand. However, as this is a single imputation method, it is unclear what the impact is on operating characteristics of the analyses (particularly type 1 error) due to not accounting for the uncertainty about the imputed values. Therefore, alternative approaches could be considered. However, if missing data occurs after an intercurrent event that is intended to be handled by the treatment policy strategy, the potential influence of the intercurrent event on the outcome needs to be appropriately reflected in the analysis (e.g., placebo multiple imputation may be reasonable after treatment discontinuation to account for potential loss of effect).

4.2.6. Use of real-world evidence

Use of real-world evidence, e.g. registries to provide pivotal evidence for an application is generally not recommended but would be acceptable as an explorative approach. The value of external controls for supporting efficacy and safety, however, is limited by methodological challenges related to the nature of the source. Observational data are usually not collected with research as their principal purpose, but may derive from different care settings, and are therefore likely to suffer from variable amounts of missing data. Due to variability in the epidemiology of the disease across regions, potential biases can arise when different databases across different regions are used. Further, the use of external controls inevitably comes with the risk of biased estimates and hypothesis tests with inflated type 1 error rates, of which the selection bias, time-related bias, or assessment bias are considered to be the most relevant amongst many other sources of bias.

4.3. Children and adolescents

As with other autoimmune diseases, the prevalence of PBC is higher in families with an affected member; 1.2% of children of patients with PBC go on to develop the disease as adults[14] [15]. However, unlike other auto-immune diseases that cause chronic liver disease, PBC is rarely diagnosed in children with only few cases of paediatric-onset PBC reported in the literature to date[16] [17] [18] [19]. As PBC is not a paediatric disease, development of new substances in the treatment of PBC in children and adolescents is not expected. Nevertheless, consultation with the paediatric committee PDCO is required.

[14] Hirschfield, GM, Dyson JK, Alexander GJM, Chapman MH, Collier J, Hübscher S, et al. The British Society of Gastroenterology/UK-PBC primary biliary cholangitis treatment and management guidelines. Gut. 2018; 67(9):1568-94.
[15] Carey EJ, Ali AH, Lindor KD. Primary biliary cirrhosis. Lancet. 2015;386(10003):1565-75
[16] Melegh B, Skuta G, Pajor L, Hegedüs G, Sumegi B. Autoantibodies against subunits of pyruvate dehydrogenase and citrate synthase in a case of paediatric biliary cirrhosis. Gut. 1998; 42(5): 753-756.
[17] Dahlan Y et al: Pediatric-onset primary biliary cirrhosis. Gastroenterol 2003; 125: 1476-1479.
[18] Kitic I, Boskovic A, Stankovic I, Prokic D. Twelve-year-old girl with primary biliary cirrhosis. Case Rep Pediatr. 2012; 2012: 937150.
[19] Liberal, R, Gaspar R, Macedo G. Pediatric-onset primary biliary cholangitis. Dig Liver Dis. 2019;51(7):1064-65.

4.4. Safety

Apart from general safety monitoring, special attention should be paid to liver safety considering the target organ and the additional risk of liver toxicity in patients with underlying liver disease. Furthermore, the underlying liver disease may hamper the evaluation of hepatic safety. The distinction of progressive disease and fluctuation and flare of the underlying liver disease from subclinical liver damage and true drug-induced liver injury (DILI) caused by an investigational agent is therefore a most important feature of the evaluation of liver safety. The identification of potential Hy's law cases, e.g. using an eDISH plot, and causality assessment are valuable parts of the evaluation of liver safety. It is recommended to include in the clinical study protocol clear rules for the management and assessment of significant liver adverse events, including potential Hy´s Law cases. A safety data monitoring committee, including one or more hepatology experts is recommended. Causality assessment of serious liver adverse events, including potential Hy´s Law cases should use expert adjudication. For the cholestatic liver diseases, the identification and assessment of potential Hy´s law cases is particularly challenging considering that elevated bilirubin in these diseases may reflect cholestasis due to the underlying disease rather than impaired liver function and thus criteria for identifying potential Hy´s Law cases may need to be adapted, e.g. by the use of INR as a marker for liver function. In certain cases, obtaining biopsies may be needed for the causality assessment[20] [21].

While patients with underlying liver disease may not be at an increased risk of developing DILI, the underlying liver disease may lead to a more severe reaction in those that develop DILI. Thus, exclusion criteria e.g. related to liver biochemistry, monitoring of liver tests during the study and stopping criteria should be defined in a way to timely identify and reduce any risk related to liver toxicity. Due to the confounding effect of the underlying liver disease, alternative approaches compared to drug development in other diseases may be needed. These alternative approaches may include different stopping rules, as well as thresholds to define clinically relevant events and the use of novel statistical approaches specifically developed for this purpose[22].

5. Primary sclerosing cholangitis

5.1. Short description of the disease

5.1.1. Epidemiology

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 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 10 per 100,000 inhabitants23. The age at diagnosis is mostly between 30 and 40 years, but children can also be affected.

The disease is frequently associated with inflammatory bowel disease (IBD), including both Crohn's disease and ulcerative colitis, which are clinically evident in 50-80% of patients with PSC[24].

[20] Kullak-Ublick GA et al: Drug-induced liver injury: recent advances in diagnosis and risk assessment. Gut 2017; 66: 1154-1164
[21] Palmer M et al: Consensus guidelines: best practices for detection, assessment and management of suspected acute drug-induced liver injury occurring during clinical trials in adults with chronic cholestatic liver disease. Aliment Pharmacol Ther 2020; 51: 90-109.
[22] 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.
[23] Molodecky NA et al: Incidence of primary sclerosing cholangitis: a systematic review and meta-analysis. Hepatology 2011; 53: 1590-1599.
[24] Karlsen TH et al: Primary sclerosing cholangitis – a comprehensive review. J Hepatol 2017; 67: 1298-1323.

5.1.2. Clinical presentation and disease course

Primary sclerosing cholangitis may be asymptomatic and patients diagnosed on the presence of cholestasis when screening at risk patients (e.g. those with IBD), or general health screening. Considerable variation in biochemistry and symptoms may be observed during the course of the disease. Symptoms usually develop with progression of the disease, but may also be present at initial

diagnosis. Pruritus is common in patients with PSC and can be disabling, leading to excoriations and decreased quality of life. Fevers, chills, night sweats, and right upper quadrant pain can also be present, which may represent episodic bacterial cholangitis from biliary obstruction rather than advanced disease. Liver biochemical tests may worsen during these episodes, but persistent jaundice usually reflects advanced disease.

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 development of the disease is generally slow. The development of end-stage liver disease may take more than 20 years[25]. Rate of disease progression may vary between one individual to the next and some patients have a more rapid course. Patients with PSC are at high risk of cholangiocarcinoma (CCA), gallbladder cancer, and colorectal carcinoma, whereas the presence of an increased risk for hepatocellular carcinoma is controversial[26] [27]. In patients with IBD, the course of the bowel disease is generally mild and the course of PSC independent of the course of the underlying IBD.

5.1.3. Diagnosis of PSC, prognosis and treatment options

The diagnosis is made based on the serum markers of cholestasis and finally on the presence of stricturing cholangiopathy on imaging, usually magnetic resonance imaging/cholangiopancreatography (MRI/MRCP). Biochemical liver tests demonstrate a cholestatic pattern, with elevation of serum alkaline phosphatase (ALP) predominating in most patients. Disproportionate elevation of ALP compared to alanine aminotransferase (ALT) and aspartate aminotransferase (AST) may occur. Serum bilirubin may also be elevated. Other tests for synthetic liver function, such as INR/prothrombin time (PT) and albumin, may be abnormal. Radiographic findings include abnormal-appearing bile ducts with wall thickening, dilations, and strictures.

A final diagnosis also requires the exclusion of secondary sclerosing cholangitis, particularly IgG4–related disease. Liver biopsy is not routinely performed but is needed in patients with suspected small duct PSC where there is involvement of small intrahepatic bile ducts only (about 10% of PSC patients) or in patients with suspected overlap with autoimmune hepatitis (between 7-14% of patients with PSC)[28] [29].

Markedly elevated levels of ALP are associated with worse prognosis but its naturally fluctuating course in PSC makes it a poor prognostic marker. There are a number of other factors that are 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" defined by endoscopic retrograde cholangiopancreatography (ERCP) are factors

[25] Boonstra K et al: Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis; Hepatology2013; 58: 2045-2055.
[26] Dyson JK et al: Primary sclerosing cholangitis: The Lancet 2018; 391:2547-2559
[27] Horley-Silva JL et al: An update on cancer risk and surveillance in primary sclerosing cholangitis. Liver Int 2017; 37: 1103-1109.
[28] European Association for the Study of the Liver. EASL clinical practice guidelines on sclerosing cholangitis. J Hepatol 2022; 77: 761-806.
[29] Lindor KD et al: ACG Clinical Guideline: Primary Sclerosing Cholangitis. Am J Gastroenterol 2015; 110: 646-659.

associated with negative outcomes (with regard to transplant-free survival[30]. Recently, the term "dominant stricture" has been proposed to be replaced by the term "relevant stricture", denoting high-grade strictures on MRCP associated with functional impairment and need for assessment of possible therapeutic interventions. A definition of "dominant stricture" on MRI is lacking. Because of this, routine use of MRI/MRCP as a prognostic tool in the management of PBC is not recommended[31].

Currently, no established medical treatment exists for the treatment of PSC. UDCA at doses up to 20 mg/kg is commonly used and is approved in some countries for the treatment of PSC, despite lack of evidence of a beneficial effect on clinical outcomes. Antibiotics are used to treat episodes of bacterial cholangitis. A therapeutic endoscopic intervention is indicated in symptomatic patients with significant extrahepatic strictures.

5.2. Patient population, study design and endpoints

According to ICH E9(R1) addendum on estimands and sensitivity analysis in clinical studies (EMA/CHMP/ICH/436221/2017), the scientific question of interest should be specified, and study features should be aligned.

5.2.1 Treatments

As no effective treatment is currently available, the acceptable comparator is placebo.

5.2.2 Patient populations

The diagnosis of PSC should be established based on the combination of elevated ALP and an abnormal cholangiography consistent with PSC, preferably by high-quality MRI/MRCP. Secondary reasons for sclerosing cholangitis need to be excluded, including IgG4 related sclerosing cholangitis. Patients with AIH overlap syndrome should be excluded from confirmatory studies in PSC.

For inclusion into studies, minimal threshold for ALP and bilirubin elevations should be defined and justified. Thresholds for inclusion into studies in PSC are not established. Therefore, the applicant is advised to seek scientific advice on the inclusion criteria. The selection of PSC patients should allow the occurrence of relevant events in the population included. The inclusion of small duct PSC patients would therefore usually require the presence of other, negative risk factors, such as ulcerative colitis. In general, patients should not have relevant recent fluctuations of serum markers, and not have relevant cholestasis at inclusion. Previous flares of cholangitis should not be an exclusion criterion, however, a relevant time-period of at least 6 months from the last flare should be defined. Concurrent antibiotic therapy should not be part of the medication at inclusion. Patients with strictures needing endoscopic/surgical intervention should be excluded. It may also be sensible to define an upper limit of other markers of liver damage and liver function (e.g. for transaminases and total bilirubin) for exclusion of patients with AIH-PSC as well as for safety reasons. The presence of CCA or gallbladder carcinoma, as well as colorectal carcinoma should be excluded to the extent possible. Patients with concomitant IBD may be included and it is recommended to stratify randomisation based on IBD (presence or absence). Usually, patients with active IBD should not be included as it may confound any assessment with regard to biochemical markers and symptoms. Similarly, concomitant medication for IBD would require to be stable for a relevant time-frame and should remain stable throughout the study duration. Relapses of IBD during the study should be managed according to standard of care and

[30] Weismüller, TJ et al. Patient age, sex, and inflammatory bowel disease phenotype associate with course of primary sclerosing cholangitis. Gastroenterology. 2017; 152: 1975–1984.
[31] Schramm C et al: Recommendations on the use of magnetic resonance imaging in PSC – A position statement from the International PSC Study Group. Hepatology 2017; 66: 1675-1688.

discontinuation of study drug should be considered. For drugs with potential efficacy in both IBD and PSC, studies should be performed separately in the IBD population and the PSC population, respectively.

For a disease modifying study, symptomatic as well as asymptomatic patients, can be included. Despite the fact that UDCA is not a recommended medication in PSC, it is in widespread use. Therefore, the inclusion of patients on concomitant UDCA can be allowed, but intake of UDCA should not be altered during the study and the randomisation should be stratified with respect to UDCA treatment at baseline.

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

5.2.3 Endpoints and summary measure

A development strategy aiming at the demonstration of effects at an early time-point using intermediate endpoints, for which a relationship to clinically relevant liver-related endpoints has at least been demonstrated by natural history studies, with a later confirmation on long-term endpoints, is an acceptable option. Considering the challenges of validating an intermediate endpoint, a regulatory strategy aiming at demonstrating efficacy based on the totality of data, with support from multiple endpoints may, however, be more realistic. Clinical endpoints are particularly important in this context for concluding on the overall benefit- risk profile.

Based on natural history data, the level of ALP – at diagnosis and after follow-up – has repeatedly been demonstrated to be associated with outcomes in PSC[32] [33]. However, there was an obvious dissociation of ALP and relevant clinical outcomes in the UDCA studies[34] [35][36].

Based on the above, ALP alone can currently not be accepted as an 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, however newer research has shown that – in addition to its obvious face validity – histology can be used to evaluate the changes in the disease status[37].

Therefore, a combined use of histology evaluation and ALP changes as co-primary endpoints are regarded to represent an acceptable intermediate endpoint for the disease for the time being, reflecting both aspects of the disease, i.e. bile duct pathology and fibrosis. Use of other methods for assessing fibrosis e.g. transient elastography or high-quality MRI will have to be justified and may become acceptable in the future depending on the degree of evidence currently being created for these methods in PSC.

Furthermore, it is suggested that a responder-type evaluation based on the criteria of therapeutic response should be the basis, defining biochemical response as a reduction of ALP to < 1.5xULN, or a combination of the reduction to < 1.5xULN with an at least 40% reduction from baseline. For the histological evaluation, based on one of the accepted staging system, a similar responder-type evaluation is proposed. The response should be defined based on an at least 1 point improvement in

[32] 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
[33] 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.
[34] Olsson R et al: High-Dose Ursodeoxycholic Acid in Primary Sclerosing Cholangitis: A 5-Year Multicenter, Randomized, Controlled Study. Gastroenterology 2009; 129: 1464-1472.
[35] Lindor KD et al: High-dose ursodeoxycholic acid for the treatment of primary sclerosing cholangitis. Hepatology 2009; 50: 808-814.
[36] Ponsioen CY et al: Design and Endpoints for Clinical Trials in Primary Sclerosing Cholangitis. Hepatology 2018; 68: 1174-1188.
[37] 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

the fibrosis stage. Stable disease (no worsening of fibrosis) could be used instead, if adequately justified. The co-primary endpoints should be supplemented with relevant secondary endpoints, including an endpoint measuring symptom improvement (e.g. pruritus, fatigue).

A later evaluation of long-term outcomes is also considered necessary for PSC, which should be done as a composite endpoint including the manifestation 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 (MRI/MRCP, transient elastography), other biomarkers (bilirubin, transaminases, but also other potential future biomarkers) as well as important clinical events in the course of the disease, such as the number of bouts of acute cholangitis, occurrence of dominant/relevant stricture, and finally the occurrence of CCA, and other malignancies. A final conclusion on the benefit-risk ratio would have to be based on the totality of these data.

Study duration

As for PBC, study duration is anticipated to be approximately 2 years, and for the demonstration of the long-term clinical outcomes a study duration of several years is anticipated.

This proposed study 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, for a long-term clinical outcome extension, an event driven evaluation will be planned for.

5.2.4 Intercurrent events and strategies

It would be expected that all intercurrent events are pre-defined in the study protocol and compared between the treatment arms. Among potential intercurrent events to be taken into account are lack of adherence to treatment (e.g. treatment discontinuation, treatment interruption, dosing deviation and intake of concomitant medication etc.), cholangitis flare with need for additional treatment with antibiotics, occurrence of malignancy and death due to other causes. 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, occurrence of malignancies, liver-related events and death should be handled with a composite strategy as treatment failure. In general, unless an alternative is justified, for other intercurrent events the "treatment-policy" strategy should be applied, i.e. the outcome regardless of the intercurrent event is of primary interest.

Similarly to PBC, estimands for secondary endpoints should be defined.

5.2.5 Study design and statistical analysis

Choices made regarding design and statistical analysis, including the handling of missing data, must be made considering the target of estimation. The same recommendations as outlined in section 4.2.5 above apply.

5.2.6 Use of real-world evidence

Same recommendation regarding the use of real-world evidence in PSC apply as outlined in the section 4.2.6 above on real world evidence in PBC above, i.e. registries to support an application would be acceptable as an explorative objective. In the context of PSC, changing incidence and -epidemiological factors that vary across regions, may impact the interpretability of some real-world data sources.

5.3. Children and adolescents

Paediatric PSC is a very rare disease, even compared to adult PSC, which is classified as orphan.

Mean age of onset of paediatric PSC is 11.2±5 years[38]. 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. Overlap syndrome of PSC and AIH is more frequent in paediatric PSC patients than in adult PSC patients[39]. Up to 30 % of all paediatric patients with PSC present with biochemical, serological and histological features typical for AIH. There is no consensus on the nomenclature of this condition (PSC–AIH overlap syndrome, PSC with features of AIH, autoimmune sclerosing cholangitis). 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[40], there is still a need to collect further natural history data to support relevant studies in paediatric patients. Similar to PSC in adults, paediatric PSC has a chronic, progressive, and relentless course. However, progression of PSC/PSC-AIH seems to be slower in paediatric patients compared to adult patients.

As in adult PSC, small duct PSC in paediatric patients appears to be a milder phenotype of the disease (compared to large duct disease) and small duct PSC appears to be more frequent in paediatric patients with PSC (14.7%) compared to adult patients with PSC (up to approximately 5%). A milder phenotype with a more favourable prognosis was also found for paediatric PSC that was associated with IBD: PSC patients with IBD had a lower rate of cirrhosis and higher survival with a native liver.

Once these natural history data are available and have been evaluated to a sufficient extent, studies in paediatric PSC may also be conducted with patients suffering from overlap conditions (especially PSC-AIH), 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). Relevant, validated biomarkers are needed for future clinical studies in paediatric PSC patients and could even serve as the basis for future paediatric prognostic calculators.

Besides the need to fully explore the PK profile in the respective population, currently no clear recommendations can be given with regard to the design of studies, and endpoints to be used. Similar to adults combined use of histology evaluation and biochemical changes are acceptable endpoints. Use of other methods for assessing fibrosis e.g. transient elastography or high-quality MRI will have to be justified and may become acceptable in the future depending on the degree of evidence currently being created for these methods in PSC. Adolescents are recommended to be included in the adult

[38] Cotter JM, Browne LP, Capocelli KE, McCoy A, Mack CL. Lack of correlation of liver tests with fibrosis stage at diagnosis in pediatric primary sclerosing cholangitis. JPGN, 2018; 66(2):227-233..
[39] Mieli-Vergani, G et al: Diagnosis and Management of Pediatric Autoimmune Liver Disease: ESPGHAN Hepatology Committee Position Statement. JPGN 2018; 66: 345–360.
[40] Deneau MR et al: The natural history of primary sclerosing cholangitis in 781 children: A multicentre, international collaboration. Hepatology 2017; 66: 518-527.

study for generating relevant data. Consultation with the agency early in the drug development (scientific advice and submission of a paediatric investigational plan (PIP)) is therefore advisable.

5.4. Safety

The recommendations for safety evaluation made above for PBC also apply for PSC. The distinction between progression of the disease and complications such as acute bacterial cholangitis and the development of CCA, make it even more challenging to identify and assess potential DILI. In addition, some additional safety issues related to PSC need to be considered. The risk of CCA, gallbladder cancer and colorectal cancer is increased in patients with PSC. In the clinical trial setting, increased monitoring may be needed, in particular with drugs with a mechanism of action that may increase the risk for any type of cancer, e.g. immunomodulators. Prior to inclusion, every effort should be made to exclude malignancy by recent imaging, endoscopy and laboratory markers, including tumour markers such as CA 19-9. Monitoring should be in place for these malignancies. Monitoring should include measuring tumour markers and imaging, e.g. annual MRI. Potential adverse effects of the drug on concomitant IBD may need to be considered (e.g. corticosteroid use).

Special attention needs to be paid to the paediatric patient population (also if included into adult studies), growth, development and sexual maturation should be followed carefully in these patients.

6. Treatment of cholestatic pruritus in PBC and PSC

PBC and PSC may both impose a significant and clinically relevant burden of symptoms on patients with the disease[41]. As such, symptom evaluation should be part of any study in these diseases and may form a basis for a claim in SmPC section 5.1 together with other evidence of efficacy.

Among symptoms associated with PBC and PSC, pruritus related to cholestasis and the retention of bile acids is common during the course of the diseases. Pruritus may become a major clinical burden dramatically impairing quality of life and even leading to suicidal ideations in the most severe cases.

Currently, no treatment is consistently approved across the EU for the symptomatic treatment of cholestatic pruritus in PBC or PSC and the approved therapies for PBC are not effective for the treatment of symptoms, including pruritus. Bile sequestrants are widely used as first-line therapy despite limited evidence. Therefore, due to this unmet medical need, it is also possible to develop new treatments which address the indication of symptomatic improvement of pruritus, without aiming generally at positively influencing the natural disease course (disease modification). Although an indication claim on symptomatic treatment is possible, it is expected that in addition special attention is put on alkaline phosphatase and bilirubin values, as well as on non-invasive markers of liver fibrosis and disease progression.

For the indication of cholestatic pruritus in PBC and PSC, separate studies in each disease should be performed. Further, it will usually be expected that effects are also evaluated in other cholestatic diseases, in order to claim a general pruritus indication associated with cholestatic diseases.

6.1. Study design, patient population and endpoints

6.1.1. Design

Despite cholestyramine being approved in some EU countries for the symptomatic treatment of cholestatic pruritus in PBC a superiority trial using placebo as comparator is acceptable.

[26] Dyson JK et al: Primary sclerosing cholangitis: The Lancet 2018; 391:2547-2559

6.1.2. Population

Patients included should have a minimum level of severity of pruritus at baseline, i.e. at least moderate symptom burden in order to allow the demonstration of a clinically relevant effect of the new drug. Randomisation should be stratified for use of anti-pruritic treatment.

6.1.3. Endpoints

Usually, a claim of efficacy should be based on an instrument measuring pruritus, supported by a more indirect evaluation of the impact of the symptom. The clinical relevance of symptomatic improvements at the chosen follow-up period should be substantiated in the future marketing authorization application.

Currently, no fully validated patient reported outcome (PRO) for pruritus in PBC and PSC exists and the development of such tools is encouraged. In the meantime, partly validated PROs can be included in studies, also for the purpose of further validation.

Other relevant secondary endpoints that should be evaluated are quality of life, sleep impairment, etc. The applicant is advised to seek scientific advice in case a new disease-specific symptom and/or health-related quality of life scale is proposed.

Bile acids are recommended to be measured as exploratory endpoints when applicable.

If the totality of the disease specific symptoms is the aim of the new treatment, it is recommended that disease specific measurements of the symptoms are part of the primary evaluation.

Study duration

Clinical studies with this restricted scope could be planned with a limited duration of placebo-controlled treatment for 3-6 months. A sufficient amount of long-term data, in order to demonstrate adequate safety should, however, also be available.

6.1.4. Intercurrent events and strategies

The evaluation of a symptomatic treatment is suggested 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, that are difficult to predict. 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.