15 March 2022
EMADOC-1700519818-746443
Executive Director

Letter of support for Sjögren's Tool for Assessing Response (STAR)

On 04 June 2021 the Applicant Assistance Publique Hopitaux De Paris requested qualification advice for Sjögren's Tool for Assessing Response (STAR) pursuant to Article 57(1)(n) of Regulation (EC) 726/2004 of the European Parliament and of the Council.

Sjögren's Tool for Assessing Response (STAR) is intended to assess treatment efficacy based on improvement of disease activity in clinical trials of drugs for primary Sjögren's syndrome (pSS).

During its meeting held on 25 – 28 October 2021, the SAWP agreed on the advice to be given to the Applicant.   During its meeting held on 08 – 11 November 2021, the CHMP adopted the advice to be given to the Applicant.

Context of Use for STAR

This advice concerns a qualification of the composite responder index STAR (Sjögren's Tool for Assessing Response) intended as primary endpoint to evaluate a treatment effect in clinical trials in primary Sjögren's syndrome (pSS).

Summary of development and validation of STAR

Combining a data-driven and formal expert-consensus approach, the Applicant has developed a candidate composite response measure (STAR) for use in the broad population of patients with pSS, regardless of disease subtype (e.g., systemic disease or not). The STAR contains 5 domains: systemic activity, symptoms, lacrimal and glandular gland function, and biomarkers of auto-immune activity. The domains are differently weighted. At the current stage there are 20 candidate versions of the STAR composite. There is one version that is preferred by the experts and put forward for the validation study: the Candidate STAR. In this candidate version a subject with a score of 5 points or more is considered a responder. See table below.

The Applicant followed a stepwise development and validation plan that has been completed up to the 'last' step, i.e., validation as (primary) endpoint in a prospective randomised controlled trial, the NECESSITY-study (NEw Clinical Endpoints in patients with primary Sjögren's Syndrome an Interventional Trial based on stratifYing patients).

The validation study proposed is a randomized double-blind controlled trial in 300 subjects with pSS. Randomisation is stratified in two cohorts, i.e., cohort 1: low systemic disease activity (ESSDAI < 5) but high level of symptoms (ESSPRI ≥5), and cohort 2: moderate-high systemic disease (ESSDAI ≥ 5) regardless of symptom level). There are 3 study arms: hydroxychloroquine-leflunomide-placebo, hydroxychloroquine-placebo-mycophenolate mofetil and triple placebo. Duration of treatment is 24 weeks with a 12-week post-treatment follow-up period.

The primary objective is to evaluate the efficacy of each active treatment combination based on proportion of responder patients according to Candidate STAR at week 24. The primary endpoint is the proportion of responders at 24 weeks; a responder is defined as a subject with a score of 5 or more on the Candidate STAR. Secondary endpoints are among others, change in ESPRI, ESSPRI responders, change in ESSDAI/clinESSDAI, ESSDAI/clinESSDAI responders.

Exploratory endpoints among others, are the proportion of responders according to alternative STAR versions.

Candidate STAR

Domain	Point	Definition of response
Systemic activity	3	Decrease of clinESSDAI ≥ 3
Patient reported outcome	3	Decrease of ESSPRI ≥ 1 point or ≥ 15% Symptoms dryness, pain and fatigue rated on 3 NRS's
Lacrimal gland function (assessed by Schirmer's test or Ocular Staining score)	1	Schirmer: If abnormal score at baseline: increase ≥ 5 mm from baseline If normal score at baseline: no change to abnormal or Ocular Staining Score: If abnormal score at baseline: decrease ≥ 2 points from baseline If normal score at baseline: no change to abnormal
Salivary gland function (assessed by unstimulated whole salivary flow or ultrasound)	1	Unstimulated Whole Salivary Flow (UWSF): If score > 0 at baseline: increase ≥ 25% from baseline If score is 0 at baseline: any increase in UWSF from baseline or Ultrasound: Decrease ≥ 25% in total Hocevar score from baseline
Biological (assessed by IgG or RF)	1	IgG: decrease ≥ 10% or Rheumatoid factor: decrease ≥ 25%
Candidate STAR responder		≥ 5 points

CHMP comments on STAR development and validation

The stepwise development and validation approach includes input from experts (health professionals and patients) using the Delphi consensus method, and quantitative analysis of historic trial data sets. This approach to item selection is principally agreed.

The Candidate STAR and its variants contain 5 domains with measurement instruments attached to those domains. The experts agreed on these, and the choices are in line with the EULAR recommendations (Ramos-Casals 2020), pointing to **face-** and **content validity**.

At a Discussion meeting with the Scientific Advice Working Party, the Applicant explained that **construct validity** will be documented using correlation analysis, discriminative ability, anchor-based analysis and magnitude of treatment effect. During the Discussion Meeting, it was also discussed that it may be problematic to accept a composite endpoint that would theoretically allow worsening in some items yet still indicate success overall. For the assessment of the overall benefit /risk of a treatment in pSS a general requirement is that none of the components should indicate/show a detrimental effect. The Applicant was encouraged to consider the possibility to include a "no worsening in the other domains" clause in the responder definition.

At the Discussion meeting the Applicant also put forward that the **test-retest reliability** of each individual item is already well-known (Seror and Bowman 2020), since the STAR development relies on measurements or assessments already performed in practice. Therefore, the reliability has been considered acceptable, and will not be reassessed. This is accepted by the CHMP.

Overall, the CHMP considers that the Candidate STAR is a potentially adequate composite responder index for assessing changes after treatment on the signs and symptoms of pSS. Still open questions relate to the performance of this outcome measure in different relevant subpopulations (e.g., according to level of systemic disease activity and disease duration) and the definition of response of each item for the Candidate STAR and its variants.

Preferably the consensus-based decision for the final STAR should be made before incorporating this endpoint in the 'pivotal' validation study. This could be done using data from currently ongoing trials that were not used in the steps up to now. Note that if, based on the results of the NECESSITY trial, another STAR variant is chosen, this is data driven, i.e., exploratory. This STAR variant will then subsequently need to be validated in a confirmatory study. It is emphasised that the purpose of the phase 4 validation exercise is not to validate the choice of domains included in STAR, but rather to validate a single, precisely defined and pre-specified STAR measure.

Overall, the validation study as proposed is largely acceptable with some remarks, that include those provided below.

Importantly, the STAR should not be designated as primary or secondary endpoint in the study where the validity of the STAR is part of the study objective/evaluation. Instead, it is recommended to include the STAR in this study as "endpoint for validation".

The usefulness of this new composite measure heavily depends on whether it is clinically interpretable. It is critical that the ≥ 5-point score, denoted as a Candidate STAR responder, corresponds to a **clinically relevant difference**. The Applicant intends to use an anchor-based method to evaluate the **meaningfulness** of this difference in the validation study. This is in principle accepted but the anchor

proposed i.e., **evaluation of change** by patient (paGA) as well as physician (phGA) is rather coarse: improvement/stable/worsening. A more fine-tuned Likert scale (5 or 7 steps) is recommended for the anchor-based method.

Finally, the treatment(s) to be tested in the prospective study ha(s/ve) been chosen based on earlier study outcomes. The one study referred to appears however very limited in size. As described, it is interpreted that at least one of the active treatment arms is expected to show a difference versus placebo. It is noted that MMF, a component of one of the investigated combination treatments, appears relatively untested (or at least with unproven efficacy) for the target indication. Lack of separation from placebo could endanger the final step of validation of the STAR index, i.e., whether the instrument is able to pick up a treatment effect.

In conclusion, the EMA acknowledges the Applicant's efforts in establishing the composite responder index STAR as a suitable primary endpoint in clinical trials that evaluate the efficacy of potential treatments in primary Sjögren's syndrome (pSS). The EMA has issued this Letter of Support to encourage the further development and validation of the STAR. A qualification opinion may be considered when the results of the final validation step are available, provided that the CHMP recommendations stated above are taken into account.

The letter of support is issued on the basis of this qualification advice.

Yours sincerely,

Emer Cooke
Executive Director