Gemeinsamer Bundesausschuss

Justification

on the Resolution of the Federal Joint Committee (G-BA) on the Finding in the Procedure of Routine Practice Data Collection and Evaluations according to Section 35a, paragraph 3b SGB V:
Valoctocogen roxaparvovec (severe haemophilia A) – submission of study protocol and statistical analysis plan

of 21 September 2023

Contents

1. Legal basis

According to Section 35a, paragraph 3b, sentence 1 SGB V, the Federal Joint Committee (G-BA) can demand the pharmaceutical company to submit routine practice data collections and evaluations for the purpose of the benefit assessment within a reasonable period of time for the following medicinal products:

1. in the case of medicinal products authorised to be placed on the market in accordance with the procedure laid down in Article 14, paragraph 8 of Regulation (EC) No 726/2004 of the European Parliament and of the Council of 31 March 2004 laying down Community procedures for the authorisation and supervision of medicinal products for human and veterinary use and establishing a European Medicines Agency (OJ L 136, 30.4.2004, p. 1), as last amended by Regulation 162 Rules of Procedure last revised: 16 December 2020 (EU) 2019/5 (OJ L 4, 7.1.2019, p. 24), or for which a marketing authorisation has been granted in accordance with Article 14-a of Regulation (EC) No 726/2004; and
2. for medicinal products approved for the treatment of rare diseases under Regulation No. 141/2000.

According to Section 35a, paragraph 3b, sentence 10 SGB V in conjunction with Chapter 5, Section 60 Rules of Procedure of the G-BA (VerfO) , the G-BA reviews the data obtained and the obligation to collect data at regular intervals, at least every eighteen months.

2. Key points of the resolution

At its session on 2 February 2023, the G-BA decided on the requirement of routine practice data collection and evaluations for the active ingredient valoctocogen roxaparvovec (Val-Rox) in accordance with Section 35a, paragraph 3b, sentence 1 SGB V.

In order to check whether the G-BA's requirements for routine practice data collection and evaluations have been implemented, the pharmaceutical company submitted drafts for a study protocol and a statistical analysis plan (SAP) to the G-BA in due time in a letter dated 30 June 2023. The documents were reviewed by the G-BA with the involvement of the Institute for Quality and Efficiency in Health Care (IQWiG).

On the basis of this review, the G-BA came to the conclusion that the requirements for routine practice data collection and evaluations in the study protocol and SAP prepared by the pharmaceutical company and submitted to the G-BA for review were insufficiently implemented.

The present declaratory resolution and the associated justification establish and justify the necessary need for adaptation of the study protocol (version 1.0 (original); 29 June 2023) and SAP (version 1.0 (original); 29 June 2023).

2.1 Necessary adjustments to study protocol and statistical analysis plan

On the necessary adjustments in detail:

1. Question according to PICO: Patient population

   According to the Val-Rox product information, there is a contraindication to treatment in active infections, either acute or uncontrolled chronic, or in patients with known significant liver fibrosis or cirrhosis. In order to ensure the positivity of the patient populations included in the routine practice data collection, the exclusion criteria should be adjusted accordingly.

2. Question according to PICO: Outcome

   The pharmaceutical company plans to start collecting the endpoints on bleeding in the Val-Rox arm five weeks after administration of the intervention, three days after the last routine factor VIII prophylaxis or 27 weeks after the last emicizumab prophylaxis, depending on which of these events occurs last. In contrast, for the comparator arm, the survey begins with the index date. This is inappropriate as the first period after administration of Val-Rox is part of the treatment strategy.

   The study protocol must therefore specify that all relevant endpoints in both study arms are collected from the index date.

3. Question according to PICO: Outcome, mortality

   The pharmaceutical company plans to collect mortality data via the safety endpoints. However, overall survival is not listed as part of the safety endpoints in table 8 of the study protocol. This is to be supplemented.

4. Question according to PICO: Outcome, bleeding

   The pharmaceutical company describes that the reason for treatment on demand should be specified as follows on the basis of the data fields in the German Haemophilia Register (DHR): Suspected bleeding, spontaneous bleeding and unknown cause. However, if the reason for treatment on demand is severe or life-threatening bleeding, it does not seem plausible to state "suspected bleeding" or "unknown reason" as the reason. In the case of severe and life-threatening bleeding, the reason for treatment on demand must be ascertained as specifically as possible, among other things by adding the data field "traumatic bleeding". The selection of "suspected bleeding" and "unknown reason" should be deleted for this case.

   The pharmaceutical company does not specify in the study protocol how severe bleeding and life-threatening bleeding are operationalised. Uniform, unambiguous and most objective definition of these events is important for a complete documentation of bleeding events with the least risk of bias. The criteria for severe bleeding and life-threatening bleeding must be precisely described in the study protocol.

According to the overall data set, the DHR currently allows the severity of bleeding to be indicated as mild, severe, life-threatening or unknown, among others. However, there are no precise criteria in the overall data set or in the DHR manual as to when bleeding should be classified as severe. There is therefore a risk that different definitions are applied in the individual study sites in general or for individual patients. If the existing data fields for bleeding events are to be used, it is necessary to provide them with clear definitions and to ensure that the data are collected according to this definition in order to carry out the routine practice data collection in the DHR.

5. Question according to PICO: Outcome, joint function

   The study protocol contains inconsistent information on the frequency of joint function assessment (1 to 2 times per year). The information in the study protocol must be standardised. An annual survey is assessed as sufficient.

6. Question according to PICO: Outcome, patient-reported outcomes (PROs) and joint function

   The pharmaceutical company describes that it assumes that patients visit every six months and plans to collect the patient-reported outcomes (PROs) and joint function at each of these visits. It therefore sets the survey at baseline and at months 6, 12, 18, 24, 30 and 36 with a tolerance range of ± 3 months for each data collection time point after baseline. This procedure is unsuitable because the survey time points are contiguous due to the choice of tolerance ranges. This would have consequences, for example, for evaluations using a mixed model for repeated measures (MMRM) in which the values are assigned to the planned points in time without taking into account the actual date of the assessment, although these may be contiguously. The study protocol should therefore define appropriate tolerance ranges for the collection of PROs and joint function that are non-contiguous. For the assessment of PROs and joint function, the pharmaceutical company could also consider an assessment by an external centre. Through this approach, the collection of PROs would not be linked to the number of visits.

   In addition, appropriate measures to avoid missing values shall be described in the study protocol.

7. Question according to PICO: Outcome, adverse events (AEs)

   According to the study protocol, only those events related to the treatment of haemophilia should be included in the analysis with regard to serious adverse events (SAEs). This does not correspond to the requirement of the G-BA, according to which all events leading to hospitalisation or death are to be collected and evaluated within the framework of the operationalisation of the endpoint of SAE. This is to be adapted.

For the specific AEs, the indication of the severity is not foreseen. This does not correspond to the requirement of the G-BA, according to which the indication of the respective severity is also required for specific AEs if indicated. In addition to recording the respective overall rates, the study protocol must therefore also specify the recording of events leading to hospitalisation or death for the specific AEs.

Also for the thromboembolic events, according to the submitted study protocol, only those events that occur in connection with the treatment of haemophilia should be included in the analyses. This is inappropriate; what is needed is an evaluation of all thromboembolic events.

In the study protocol, it is planned to collect the specific AE "severe liver damage", operationalised as liver failure or cirrhosis. In the statistical analysis plan, the operationalisation is specified to the effect that liver failure, liver fibrosis and liver cirrhosis are surveyed on the basis of Child-Pugh criteria A, B and C. The operationalisation of the specific AE "severe liver damage" is to be reviewed to ensure that symptomatic liver damage that is not based on laboratory parameters alone is recorded.

The pharmaceutical company describes that for the specific AE "malignant neoplasms" there may be differences in data collection practice between the study sites due to the current collection as free text in the DHR. For the implementation of the routine practice data collection in the DHR, it must be ensured that the specific AE is documented in a uniform manner. Appropriate measures or definitions in this regard are to be outlined in the study protocol.

The pharmaceutical company plans to censor the patients after a change of treatment for the evaluation of the AEs. In the view of the G-BA, the need for renewed factor VIII therapy after treatment with Val-Rox does not constitute a classic change of treatment. The gene therapy mode of action of Val-Rox means that treatment cannot be discontinued after a single dose, as a functional copy of the factor VIII gene is transfected into specific liver cells by the adeno-associated viruses (AAV5) and subsequently translated in these cells. In addition, patients have usually already received factor VIII therapy over a longer period of time prior to treatment with Val-Rox. Therefore, for the assessment of the long-term benefits and harms of Val-Rox, the side effects in the intervention arm, which may arise if renewed factor VIII treatment is required, are also considered relevant. For the patients in the intervention arm, it is therefore necessary to collect the AEs until the end of the study. Patients who switch to treatment with Val-Rox during the course of the study and who have a sufficiently long observation period in relation to the gene therapy will be assigned to the intervention arm. The AEs assessment must also be carried out for these patients until the end of the study.

Patients who switch to treatment with Val-Rox during the course of the study and do not have a sufficiently long observation period in relation to the gene therapy should be assigned to the comparator arm. As part of the regulatory obligations to implement post-authorisation measures, further long-term safety data must also be collected for Val-Rox. Therefore, even for those patients who switch to treatment with Val-Rox late in the comparator arm relative to the duration of observation, it is considered proportionate to collect AEs until the end of the study, thus avoiding unequal treatment between the study arms.

In the conclusion, it must be specified in the study protocol that the endpoints on side effects are collected in both study arms until the end of the study. Regarding the evaluation of the AEs, please refer to the explanations below on the evaluation of the AE data (paragraph 31).

8. Question according to PICO: Outcome, supplementary information

   The endpoint "time of resumption of prophylactic treatment" is operationalised by the pharmaceutical company in the Val-Rox arm as four doses of emicizumab or 18 doses of factor preparation with extended half-life or 24 doses of factor preparation with normal half-life. This does not correspond to the requirement of the G-BA, according to which the resumption of any prophylactic treatment must be documented and presented in the benefit assessment.

9. Data source/ study design: General

   In the study protocol and SAP, the pharmaceutical company describes at various points that it only wants to determine definitions, operationalisations (e.g. of endpoints or confounders) or evaluations on the basis of the observed data or in the course of the routine practice data collection. This approach is unsuitable for conducting the routine practice data collection. All relevant data to be collected and associated evaluations must be defined a priori and described in the study protocol and SAP before the start of the routine practice data collection.

10. Data source: Collection of baseline data

    The pharmaceutical company intends to assess the inclusion and exclusion criteria on the basis of the patient characteristics documented when the patient was admitted to the DHR. This procedure is inappropriate as it only allows patients to be recruited into the routine practice data collection for whom consent to individual reporting in the DHR has already been obtained. In addition, it cannot be ensured that the patient characteristics are up to date on the index date (e.g. liver status, inhibitors).

    It must therefore be specified in the study protocol that all inclusion and exclusion criteria with clear operationalisation are collected on the index date.

For patients who switch to treatment with Val-Rox in the course of the study and are assigned to the intervention arm due to a sufficiently long observation period under gene therapy, renewed survey of baseline characteristics at the time of Val-Rox application is required to ensure a proper evaluation of these patients in the intervention arm.

The pharmaceutical company also describes in the study protocol that it only intends to access patients for whom data have already been documented in the DHR for at least 12 months in order to collect the baseline data. This would mean that the routine practice data collection would only include patients for whom consent to individual reporting had already been given 12 months before inclusion in the data collection. This restriction is considered inappropriate to ensure sufficient recruitment of patients with individual reporting for the routine practice data collection and should therefore be deleted.

The pharmaceutical company plans to present the age in 5-year categories. This is improper. The study protocol must specify that the age of the patients is clearly collected.

For the planned sensitivity analyses on AAV5 status, the virus type against which antibodies are present must be recorded in the data source used. In addition, it must be ensured that any AAV5 status collected is recorded for the patients of both study arms. This is to be specified in the study protocol.

11. Data source: Definitions and operationalisation

The pharmaceutical company describes in the study protocol that it intends to refine the definition of the variables to be collected based on the raw data of the routine practice data collection in the course of the study. This procedure is unsuitable for the implementation of the routine practice data collection, since, as already mentioned under paragraph 9), all data to be collected must be predefined with their corresponding operationalisation and described in the study protocol or SAP.

For operationalisation, standard classifications are to be used where possible to ensure uniform coding of the data. The pharmaceutical company describes that, for example, information on comorbidities is recorded as free text and notes that this can lead to data on comorbidities not being meaningfully evaluated.

It is the responsibility of the pharmaceutical company to determine the operationalisation and definition of the data to be collected in such a way that usable data from the routine practice data collection are available for all relevant patient characteristics and confounders as well as endpoints.

Currently, the data fields described in SAP are not yet established in the DHR. The collection of the required data must be ensured in the selected data source before the start of the routine practice data collection.

12. Data source: Confounders

The pharmaceutical company describes in Annex 2 to the study protocol that it identifies relevant confounders based on the patient characteristics of the 1-arm observational

study 270-902 and the 1-arm study 270-301. It involves clinical experts for the selection of potentially relevant confounders. The pharmaceutical company does not plan to conduct a systematic literature search until the interim analysis after 18 months. The pharmaceutical company then plans to match the relevant confounders identified with the variables collected in the DHR. In addition, the pharmaceutical company plans to focus its literature research primarily on data collected in Germany.

The pharmaceutical company's approach does not ensure that a sufficiently complete list of potentially relevant confounders is identified at the start of the study. It is not known how the 14 potential confounders for the two studies 270-902 and 270-301 were identified and whether they were based on a systematic search involving experts.

The pharmaceutical company selects from the list of 14 potential confounders only those that show a statistically significant association with the log-transformed annualised bleeding rate (ABR) in the population of the study 270-902. This procedure is inappropriate for a confounder adjustment as it requires adjustment for prognostic factors under the control as well as under the intervention, and for effect modifiers.

The list of potential confounders includes, for example, the region, which, according to current planning, has no relevance for the present routine practice data collection, since the DHR, which focuses on Germany, is used as the primary data source and thus, different regions are not available. In addition, there are differences in the confounders identified by the DHR and the pharmaceutical company in Annex 2 to the study protocol. The DHR has identified age, body weight, bleeding duration and frequency, reason for therapy, underlying mutation, comorbidities, family history and medically relevant events as potential confounders. The pharmaceutical company specifies age, BMI, region, number of affected joints, annual bleeding rate at baseline, factor VIII consumption at baseline and the type of factor VIII preparations used. In addition, the pharmaceutical company notes in the study protocol under the section of limitations of the observational study that individual lifestyles of the patients, for example physical activity, cannot be mapped and this can lead to risk of bias.

The current selection of confounders is not sufficiently comprehensible and is considered inappropriate.

The pharmaceutical company's planned procedure of conducting the systematic literature search only for the interim analysis after 18 months and then retrospectively comparing whether the DHR has collected data on the identified confounders is unsuitable for conducting the routine practice data collection. Due to the planned approach, there is a risk that data on important confounders will not be properly collected in the selected data source, thus not being adequately considered in the analysis.

It should also be noted that focusing the literature review on data collected in Germany is inappropriate.

In conclusion, a systematic literature search for potentially relevant confounders should be carried out before the start of the routine practice data collection and supplemented

with the involvement of experts. The section on the identification and definition of confounders in the study protocol must be thoroughly revised. It must be ensured that all relevant confounders identified a priori in the selected data source are collected appropriately from the beginning of the routine practice data collection. If it is not possible to collect certain parameters, this uncertainty must be addressed in the study protocol and its consideration in the interpretation of the results must be described.

13. Data source: Reporting dates

    For the DHR, there is an obligation to report 1-time a year. The pharmaceutical company links the data collection to the regular patient visits and specifies in the study protocol that the documentation in the DHR should take place at least once a year, but can take place more frequently if this follows the respective visit.

    The reporting dates in the selected data source must ensure that the data from the routine practice data collection are available for timely submission of the interim analyses specified in the resolution of 2 February 2023 and of the dossier for the new benefit assessment. This requirement must be saved in the study documents.

14. Data source: Completeness of the data

    For the implementation of the routine practice data collection in the data source selected by the pharmaceutical company (DHR), adjustments are necessary with regard to the data fields to be collected, definitions and operationalisations, among other things. In order to ensure that the data collection is appropriate and as complete as possible, the study protocol must describe measures that are taken for the training of the treating physicians with regard to the collection of the data relevant for the routine practice data collection.

    In addition, the study protocol must address measures that are carried out to ensure the completeness of the data for each patient in order to keep missing values low.

    In order to achieve the most complete data collection possible, it is necessary from the perspective of the G-BA that, within the framework of the selected data source, all relevant data for the routine practice data collection are obligatory and not only optional data fields for data entry. Accordingly, the resolution on requirements of 2 February 2023 stipulated as a requirement for the data source that specifications must exist to ensure the completeness of the data collection time point and the completeness of the data collection time points. The obligatory assessment of the data fields relevant for the routine practice data collection must be ensured by the selected data source.

15. Data source: Source Data Verification

    The pharmaceutical company describes that it intends to perform Source Data Verification (SDV) for an estimated 100% of the data fields for the inclusion and exclusion criteria and the primary endpoint by matching the data entered into the DHR at the study sites with the aggregated data submitted by the DHR. An SDV involves the matching of the source file (usually the patient record) and the entries made in the register. A comparison should be made on site between the data reported to the DHR and the

corresponding information in the respective patient record. The pharmaceutical company refers to the Study Monitoring Plan (SMP) for further information, which is not available. Thus, it was not possible to verify how the SDV for the routine practice data collection is carried out.

The technical process of source data verification must be described clearly and comprehensibly in the study protocol. The study monitoring plan must be attached to the study protocol or submitted separately for the re-examination of the study documents.

16. Study design: Estimand

    The pharmaceutical company determines the hypothetical estimand as the primary estimand of the study without sufficiently specifying the estimand. However, the primary estimand of routine practice data collection has to comply with the treatment policy strategy and includes in particular the evaluation according to the ITT principle for all patient-relevant endpoints. The pharmaceutical company must therefore adapt both the study protocol and the SAP accordingly; this applies in particular to the points concerning the index date and the start of the follow-up as well as the handling of treatment changers.

17. Study design: Recruitment

    The study protocol does not yet describe any measures to bring about the transfer of a higher percentage of patients from collective reporting to individual reporting. In order to ensure the recruitment of the required sample size for the present routine practice data collection, measures to increase the percentage of patients with individual reporting are to be described in the study protocol.

18. Study design: Assignment to the treatment groups

    The pharmaceutical company does not provide any information in the study protocol or in the SAP on how to deal with cases where patients were initially assigned to the comparator arm during the recruitment period and switch to treatment with Val-Rox during the course of the study.

    In the present situation, a strategy in the sense of an ITT evaluation is to be pursued, whereby patients who switch from the comparator therapy to Val-Rox in the course of the observation are assigned to the study arms, depending on the observation period under the comparator therapy. Those patients who have already been observed for an appropriately long time under treatment with factor VIII preparations so that meaningful data are already available for the comparator group (e.g. 2 years with a planned observation period of 3 years) and only then switch to Val-Rox are to be evaluated in the comparator arm and further observed until the end of the study. Patients who switch to treatment with Val-Rox after a shorter period of time and for whom an adequate observation period under gene therapy can still be expected are to be evaluated in the intervention arm. For these patients, the time of switching represents the observation

start for the routine practice data collection, requiring new baseline data collection (see also paragraph 10). The observation period under comparator therapy of these patients is not to be considered for the routine practice data collection.

For this procedure, the study protocol and statistical analysis plan shall specify how the above-mentioned adequate observation period is defined. According to the resolution on requirements of 2 February 2023, patients should be observed for at least 3 years. It is also important to consider how to deal with patients who switch to gene therapy shortly after the scheduled recruitment period, as these patients can also be expected to have an appropriate observation period under gene therapy.

19. Study design: Sample size planning

When planning the sample size, it must be ensured in accordance with the resolution on requirements of 2 February 2023 that a sufficient number of patients undergoing factor VIII therapy are recruited for the routine practice data collection. The estimated sample size in this case must refer to patients on factor VIII therapy.

The information on sample size planning is inconsistent between the study protocol and the SAP. Contrary to the information in the SAP, the pharmaceutical company describes in the study protocol that the underlying test for sample size estimate corresponds to a t-test. Since a dependency structure of recurrent events is assumed for the primary endpoint, the information on the methodology of sample size estimate in the study protocol is incorrect. The section in the study protocol is to be corrected by the pharmaceutical company according to the information in the SAP.

In addition, the measures taken to recruit a sufficiently high number of patients under treatment with factor VIII preparations into the routine practice data collection shall be described.

20. Study design: Discontinuation of study participation

The pharmaceutical company states that the treating physician may terminate the study participation of patients on the basis of their clinical assessment. The patients in question would thus be excluded from further observation, which can lead to a clear bias. This approach is unsuitable for conducting the routine practice data collection. The criteria used to remove patients from the study and thus, from observation must be clearly defined and established a priori.

21. Study design: Discontinuation criteria

According to the resolution on requirements of 2 February 2023, information on discontinuation criteria due to futility must be presented in the study protocol and SAP. There is insufficient information on this in the documents submitted by the pharmaceutical company. The exact discontinuation criteria are to be added. It should be noted that any decision to discontinue data collection and to change the sample size estimate must be made in consultation with the G-BA.

22. Study design: Interim analyses

By resolution of 2 February 2023, the G-BA requires evaluations on interim analyses 18, 36 and 54 months after the start of the routine practice data collection. The pharmaceutical company also plans an interim analysis after 6 months. For each of the interim analyses, the pharmaceutical company also plans to perform a test for discontinuation due to futility. The rationale for conducting an additional interim analysis after 6 months and the associated futility test is not apparent from the information in the study protocol and SAP.

The planned interim analysis 6 months after the start of the routine practice data collection is to be deleted. The interim analyses are to be carried out according to the time points stored in the resolution of 2 February 2023 18 months, 36 months and 54 months after the start of the routine practice data collection. Irrespective of this, the pharmaceutical company must submit information on the course of the data collection to the G-BA 6 months after the start of the routine practice data collection.

23. Evaluation of the data: shifted hypothesis boundary

The study documents show that the hypothesis boundary used to test primary and secondary endpoints is not shifted.

The effect to be assumed between Val-Rox and the comparator is composed of the true difference between the two treatment options and the bias due to the non-randomised study design. Due to unknown confounders, a statement on the benefit or harm of an intervention can only be derived from a certain effect magnitude. The specific threshold results from the quality of the data.

For the evaluation of the data obtained, it must therefore be stipulated that a shifted hypothesis boundary of 0.2 to 0.5 is taken into account depending on the quality of the data collection and evaluation.

In addition, the significance of the data collected in the context of the routine practice data collection is determined by the quality of the data in the specific case, for example, by the knowledge of relevant confounders. Therefore, a section should be added to the study protocol and SAP that addresses the interpretation of the results of the data, taking into account the non-randomised study design and using an appropriate shifted hypothesis boundary (in the range between 0.2 and 0.5).

24. Data evaluation: Evaluation population

The pharmaceutical company describes in the SAP that the evaluation population may differ from the total study population, as it only wants to include patients in the analyses for whom complete data are available on all variables that are included in the propensity score. This procedure is unsuitable for the evaluation of the data from the routine practice data collection, as there is a high risk that a relevant percentage of patients from the total population will not be included in the evaluation population. This requirement should therefore be deleted. Instead, measures must be established to ensure the completeness

of data for each patient on the one hand and to define clear criteria on how to deal with missing values in the analyses on the other.

25. Data evaluation: Sensitivity analyses

In order to investigate the influence of gene therapy in the comparator group, for the evaluation of the endpoints in the categories of mortality, morbidity and health-related quality of life, sensitivity analyses are to be pre-specified in which patients who switched to Val-Rox in the comparator arm and are not assigned to the intervention group are censored at the time of switching.

The literature describes procedures for dealing with the situation when a new therapy is not started in both treatment groups at the start of observation, such as the prevalent new user design. The aim of these procedures is to reduce any bias caused by an incorrect choice of observation start. Statistical methods for evaluating data from this study design take into account data prior to the index date. Methods for statistical evaluation are described in the literature, so that the evaluations are to be pre-specified as sensitivity analyses. It should be noted that potential time-dependent confounders must be collected continuously during the study.

In the resolution on requirements of 2 February 2023, the G-BA defines a therapy according to doctor's instructions as a comparator for the present routine practice data collection, taking into account recombinant or human plasma-derived blood coagulation factor VIII preparations and emicizumab. Accordingly, data on treatment with recombinant or human plasma-derived blood coagulation factor VIII preparations and emicizumab will be collected for the routine practice data collection. The pharmaceutical company describes that it intends to compare the different therapies within the comparator arm with Val-Rox. An analysis comparing only factor VIII preparations or emicizumab against Val-Rox is not described. In the view of the G-BA, sensitivity analyses for the separate evaluation of the data on Val-Rox versus the data on factor XIII preparations and on emicizumab are necessary for the evaluation of the routine practice data and must be pre-specified accordingly in the study protocol and SAP.

26. Data evaluation: Subgroup analyses

The pharmaceutical company does not describe whether and how statistical tests for subgroup differences are performed. The statistical tests for subgroup analyses shall be described in the SAP. For the benefit assessment according to the specifications in the dossier submissions, it should also be examined whether subgroup analyses should be conducted for further relevant characteristics. This applies in particular to analyses on the characteristics of disease severity and age.

Furthermore, the different specifications regarding the consideration of the AAV5 status characteristic between SAP (subgroup analysis) and study protocol (sensitivity analysis) must be standardised.

27. Data evaluation: Propensity score method

Due to the expected small number of patients in the intervention arm, the pharmaceutical company plans to consider only six confounders in the estimation of the propensity score (PS). In the view of the G-BA, this approach is unsuitable as all important confounders must be taken into account in the regression model for estimating the PS to achieve the balance for all important confounders. In the event that the regression model does not converge using all important confounders, an appropriate alternative strategy shall be recorded in the SAP. The results have to be discussed in relation to the methodology used, and the interpretation has to take into account the balance of the variables not included in the model.

As the main analysis for confounder adjustment, the pharmaceutical company chooses the weighting procedure of Standardised Mortality Ratio Weighting (SMRW), whereby patients who were treated with Val-Rox each receive a weighting of 1. Patients in the comparator arm are weighted so that their confounder distribution corresponds to the population in the intervention arm. The pharmaceutical company does not provide sufficient information on the variance estimator to be used. The SMR weighting procedure is inappropriate for routine practice data collection as it refers to the Average Treatment Effect (ATE), whereas the SMR weighting procedure leads to an Average Treatment Effect in Treated [ATT] estimator. The ATE can be estimated, for example, using Inverse Probability of Treatment Weighting (IPTW). The study protocol and the SAP shall be adapted accordingly.

The handling of extreme weights is not clearly described by the pharmaceutical company. For the identification of extreme values, it refers to its procedure for identifying extreme values in observed baseline characteristics and endpoints in the SAP. If extreme weights are observed, it is planned to examine the characteristics of the patients included in the PS estimate for plausibility. The specifications for handling extreme weights must be revised and the handling of extreme weights must be specified in a comprehensible and unambiguous manner.

For the assessment of overlap, there is only the indication that the histograms and density curves of the PS are visually examined and the c-statistics are reported. However, it is not defined when these examinations lead to sufficient overlap. This is to be supplemented.

The pharmaceutical company specifies a multi-step procedure in the SAP to deal with imbalance. This includes the transformations of the output variables and the addition of interaction terms. However, the pharmaceutical company does not describe the exact methodology for dealing with imbalances, nor does it cite appropriate literature. Nor does it discuss why this approach can adequately correct for imbalance in the present situation. Furthermore, the pharmaceutical company does not specify the criteria according to which the model selection for the final analyses is made. These aspects are to be added in the SAP.

According to the resolution on requirements of 2 February 2023, the pharmaceutical company should define a clear hierarchy of eligible PS methods and the test criteria for selecting the most robust methods. The presentation of the test criteria must contain specific information on the verification of sufficient overlap and balance. The information provided does not meet this requirement. The pharmaceutical company does not describe a hierarchy of PS procedures, nor is there sufficient information on overlap and balance testing. The specification of a main analysis by means of weighting and a sensitivity analysis with matching does not correspond to a hierarchical procedure. In addition, ambiguities also arise in the description of the procedure, e.g. with regard to the matching ratio for the sensitivity analyses and the meaningfulness of the defined main and sensitivity analysis against the background of the expected distribution of patients between the intervention and control arms in a 1:5 ratio. Against this background, a clear hierarchy of eligible propensity score methods and the test criteria for selecting the most robust method must be defined in the SAP. In this context, specific information on the verification of sufficient overlap and balance shall be provided.

Furthermore, the pharmaceutical company plans to present a naive analysis, whereby the decision algorithm remains unclear and no consequences of a naive analysis for the interpretation of results are described. It should be added to the SAP which consequences result if no propensity score procedure can be identified with which a sufficient overlap and balance can be achieved.

The pharmaceutical company describes in the SAP that it estimates the PS for the primary endpoint but not for other secondary endpoints. This section is not understandable because the PS represents the probability of a patient receiving the intervention, depending on the distribution among the confounders. Thus, the estimation of PS is independent of the endpoints considered. This must be corrected in the SAP.

The pharmaceutical company states that it will perform sensitivity analyses for the primary endpoint. Sensitivity analyses for additional endpoints are only performed, depending on the difference in the analyses of the primary endpoint. For the evaluation of the routine practice data, it is necessary to check the robustness of all patient-relevant endpoints. Therefore, sensitivity analyses with different propensity score methods have to be performed for all patient-relevant endpoints and not only for the primary endpoint.

After successful application of a PS procedure, it should be carefully checked whether the patient population resulting from the PS procedure corresponds to the original target population of the routine practice data collection to a sufficient extent. If this is not the case, the sub-population of the original target population to which the analyses resulting from the PS procedure refer shall be described. In PS matching in particular, patients who meet the inclusion criteria are excluded as expected during the analysis. Therefore, the necessity of a detailed description of the patient population resulting from the application of the respective PS procedure, including the necessity of a comparison of this patient population with the original target population of the routine practice data collection in the SAP must be compulsorily added.

28. Evaluation of the data: binary endpoints

    For binary endpoints, in the view of the G-BA, the relative risk is preferable as an effect measure for comparable durations of observation. Therefore, for binary endpoints with a comparable duration of observation, the relative risk should be defined as the effect estimator.

29. Data evaluation: Bleeding

    The pharmaceutical company plans to evaluate the bleeding endpoint, operationalised via the ABR, using a weighted negative binomial model. The information provided by the pharmaceutical company is incomplete. The evaluation should consider all events observed between the index date and the end of the study. Due to individual observation periods, the specification of an offset in the model is necessary.

    The pharmaceutical company further describes a zero-inflated negative binomial model. However, the application of this model in SAP remains unclear. On the one hand, it is described that the model is used when there are "too many" patients without bleeding events. The pharmaceutical company does not specify the limit at which it assumes that there are too many patients. On the other hand, the pharmaceutical company describes that this model is used when the primary analysis (weighted negative binomial model) does not converge. A comparison of the Akaike Information Criterion (AIC) of the 2 models helps in deciding the result to be used. This procedure is incomprehensible as no valid results are available from the primary model. The application of the zero-inflated negative binomial model shall be clearly specified.

30. Data evaluation: patient-reported outcomes (PROs)

    The pharmaceutical company plans to evaluate the Haemo-QoL-A and the BPI-SF as continuous data. It plans to do this using generalised mixed models for repeated measures. However, these models need to be described in detail, especially the effect measure. In addition, the pharmaceutical company indicates considering responder analyses for improvement or deterioration. According to IQWiG's methods paper, responder analyses are preferred for patient-reported endpoints. For this purpose, a response criterion corresponding to at least 15% of the scale range of the respective instrument must be predefined. For the interpretation of the result of the responder analysis, information on the course of the study, e.g. in the form of course curves, is required.

31. Evaluation of the data: adverse events (AE)

    For all AE endpoints, it is planned to comparatively analyse the respective endpoints only if at least 10 events occur in both study arms. This procedure is unsuitable for the evaluation of the routine practice data, as all patient-relevant endpoints must be submitted for the benefit assessment, regardless of the number of events that occurred. Furthermore, the planned evaluation of the AE is not fully described in SAP and needs to be specified. In order to map lower and higher damage, a comparison of the results of

both treatment arms, e.g. based on the relative risk, is necessary. In the case that very few events or no events are observed in a treatment arm, appropriate statistical procedures such as the Firth correction and profile likelihood confidence intervals should be used.

As stated under paragraph 7, the pharmaceutical company plans to censor the patients after a change of treatment for the evaluation of the AE. The G-BA does not consider the renewed need for treatment with factor VIII preparations in the intervention arm to be a classic change in treatment, so that no censoring of patients in the intervention arm has to take place if factor VIII therapy is resumed.

For patients who switched to treatment with Val-Rox in the comparator arm and continue to be assigned to the comparator arm, an analysis in which these patients are not censored at the time of switching as well as another analysis in which these patients are censored at the time of switching shall be defined for the evaluation.

32. Data evaluation: Dealing with missing values

    According to the resolution on requirements of 2 February 2023, information should be provided on the expected scope and reasons for missing data, as well as measures to avoid missing data and evaluation strategies to deal with missing data.

    Regarding the possible replacement of missing values, the pharmaceutical company provides contradictory information in the SAP in the section on PS estimation (section 7.8.1) and the section on missing values (section 7.9.5.4). Meaningful replacement strategies for missing data shall be outlined and the corresponding methodology shall be pre-specified.

    The planned replacement of the month potentially leads to significant risks of bias and is not appropriate. Appropriate efforts are rather required to minimise the percentage of missing values in date specifications. Corresponding explanations are to be added to the study documents.

In order to avoid inconsistencies, the pharmaceutical company must check whether the need for changes in the study protocol described here leads to corresponding subsequent changes in the SAP and vice versa.

In addition to the mandatory adaptations, the G-BA makes the following recommendations for further adaptations of the study protocol and the SAP:

1. Question according to PICO: Inclusion criteria

   It is recommended that the inclusion criteria not be restricted to subjects who have already been treated with factor VIII preparations for 12 months. Since it can be assumed that all patients in the present therapeutic indication receive factor VIII prophylaxis, the restriction has no direct consequence. However, no justification for this inclusion criterion emerges from the study documents. If the restriction is maintained by the pharmaceutical company, a justification should be added to the study documents. Irrespective of this,

reference is made to the mandatory requirements regarding the collection of baseline data.

2. Data source: Reporting dates

   It is recommended that the documentation of the collected data be carried out uniformly for all patients directly after the respective visit, if possible, in order to avoid reporting delays, and that this be specified accordingly in the study documents.

3. Data evaluation: Propensity score method

   When revising the envisaged procedures for confounder adjustment, it should be taken into account that the procedure envisaged so far for dealing with extreme weights leads to trimming, regardless of the observation of extreme weights. The procedure for dealing with extreme weights should therefore be revised in this respect, depending on the specific methodology chosen.

2.2 Deadline for submission of the revised study protocol and statistical analysis plan

The revised study protocol and the revised SAP are to be submitted to the G-BA by 19 October 2023.

When submitting the revised version of the SAP and the study protocol, the pharmaceutical company must ensure that the changes made can be completely and clearly understood. For this purpose, a version of the documents must usually be submitted in which the changes have been marked in detail, as well as a current version of the documents without marking the changes. Amendments that do not result from the need for adjustment set out in this resolution and the justification shall be justified separately.

3. Process sequence

In order to check whether the requirements of the G-BA for routine data collection and evaluations for the active ingredient valoctocogen roxaparvovec have been implemented as specified in the resolution of 2 February 2023, the pharmaceutical company submitted drafts of a study protocol and a SAP to the G-BA. The documents were reviewed by the G-BA with the involvement of IQWiG.

The issue was discussed in the working group WG RPDC and in the Subcommittee on Medicinal Products.

At its session on 21 September 2023, the plenum decided on the result of the review regarding the submitted study protocol (version 1.0 (original); 29 June 2023) and the statistical analysis plan (version 1.0 (original); 29 June 2023).

Chronological course of consultation

Session	Date	Subject of consultation
WG RPDC	21 August 2023 7 September 2023	Consultation on the study protocol and statistical analysis plan (SAP)
Subcommittee Medicinal products	11 September 2023	Consultation on the result of the review of the study protocol and SAP
Plenum	21 September 2023	Resolution on the result of the review of the study protocol and SAP

Berlin, 21 September 2023

Federal Joint Committee (G-BA)
in accordance with Section 91 SGB V
The Chair

Prof. Hecken