lDoc. Ref. No.: CSR AEZS-130-052_01
Date of Issue: 16-June-2017

Clinical Study Report

Excerpt from final Version 1, dated 16-Jun-2017:

Section 2 Synopsis

Confirmatory Validation of Oral Macimorelin as a Growth Hormone (GH) Stimulation Test (ST) for the Diagnosis of Adult Growth Hormone Deficiency (AGHD) in Comparison with the Insulin Tolerance Test (ITT)

(Short title: Validation of Macimorelin as a Test for Adult Growth Hormone Deficiency)

Clinical Phase	**III**
Design	**Open-label, randomized, single dose, 2-way crossover**
Clinical Study No.	**AEZS-130-052**
EudraCT No.	**2015-002337-22**
Study Protocol Version (Date)	**2.0 (18-Feb-2016)**
Study Period	**03-Dec-2015 – 28-Nov-2016**
Coordinating Investigator	**Jose M. Garcia, MD, PhD**

This study was performed in compliance with applicable Good Clinical Practices (GCP) and regulations, including the archiving of essential documents.

Sponsor **Aeterna Zentaris GmbH**
Weismüllerstraße 50
D-60314 Frankfurt am Main
Germany

2 Synopsis

Name of Sponsor/Company: Aeterna Zentaris	Volume: Page:	(For National Authority Use Only)
Name of Finished Product:		
Name of Active Ingredient: Macimorelin		

Title of study:

Confirmatory validation of oral macimorelin as a growth hormone (GH) stimulation test (ST) for the diagnosis of adult growth hormone deficiency (AGHD) in comparison with the insulin tolerance test (ITT). (Short title: Validation of Macimorelin as a TEST for Adult Growth Hormone Deficiency).

Coordinating investigator:

Jose M. Garcia, MD, PhD; Associate Professor, Department of Medicine; Division Gerontology & Geriatric Medicine; GRECC, Puget Sound VA HCS and University of Washington School of Medicine; 1660 South Columbian Way; Seattle, WA 98108-1597, U.S.; Jg77@uw.edu; phone: +1 206 764-2984.

Study centers:

30 sites in 9 countries; 25 sites in Europe (Austria, Germany, Spain, France, Italy, Poland, Serbia, and UK) and 5 sites in the USA.

Publication (reference):

None

Study period: First subject randomized: 03-Dec-2015, Last subject completed: 29-Nov-2016	Clinical phase: III

Objectives:

Primary: To validate the use of single dose oral macimorelin for the diagnosis of AGHD ('macimorelin GHST'), using the insulin tolerance test as comparator GHST.

Secondary: To characterize the safety profile of single dose oral macimorelin in suspected AGHD subjects.

Methodology:

Open-label, randomized, multicenter, multinational, 2-way crossover study.

Study subjects were assigned to groups of descending likelihood of having AGHD:

Group A: High likelihood of growth hormone deficiency (GHD):

- Structural hypothalamic or pituitary lesions and low insulin-like growth factor 1 (IGF-1), and/or
- Three or more pituitary hormone deficiencies (PHD) and low IGF-1, or
- Childhood onset GHD with structural lesions and low IGF-1.

Group B: Intermediate likelihood of GHD:

- Eligible subjects not qualifying for either high or low likelihood (Group A/C)

Group C: Low likelihood of GHD:

- One risk factor for GHD only, such as history of distant traumatic brain injury (TBI) or one PHD only with otherwise normal pituitary function or
- Isolated idiopathic childhood onset GHD without additional pituitary deficits.

Group D: Healthy control. Healthy subjects matching Group A subjects by sex, age, body mass index (BMI), and estrogen status (females only).

The sequential order of the GHSTs for the suspected AGHD subjects was determined by stratified randomization by Group; healthy control subjects (Group D) were tested in the same sequence as the matched Group A subjects. The macimorelin-GHST ('MAC') and the ITT of the core study were to be performed 7 days to 1 month apart.

Serum concentrations of growth hormone were measured at pre-defined time points before and after GHST administration of macimorelin or insulin.

Timepoints for GH and PK measurement for the MAC (Investigational GHST): Pre-dose, 30, 45, 60, 90 minutes after oral administration of macimorelin.

Timepoints for GH measurement for the ITT: Pre-dose, 15, 30, 45, 60, 90, and 120 minutes after insulin injection.

Serum concentrations of GH for a given subject were provided by the central laboratory to the investigator only after the appropriateness of the test performance had been adjudicated by a Data Review Committee (DRC) comprising two clinical investigators and one sponsor representative, and a GHST with critical deviations could have been repeated. In this way, the blinded status of the investigator and the DRC as to the outcome of the stimulation test was maintained until both GHSTs of the core study were completed.

The following cut-off values for stimulated GH levels were used:

- MAC: GH: 2.8 ng/mL
- ITT: GH: 5.1 ng/mL

A peak GH value below the cut-off value (i.e., < 2.8 or < 5.1, respectively) was considered 'test positive'.

The ITT was considered as comparator (non-reference standard) to assess positive and negative agreement of both GHSTs, based on the above cut-off points.

Repeatability extension (Amendment no 1): a second macimorelin-GHST (Test 3) was performed in a subset of subjects that had completed the study procedures according to the core study. The amendment was implemented in selected study sites in Europe only, and its objective was to generate repeatability data rather than hypothesis testing. Peak GH levels were to be compared following repeated treatments with macimorelin. At or after the End-of-Study visit of the core study, 30 patients, were to be enrolled: i.e., the first 10 patients each from the low, intermediate, and high likelihood AGHD Groups (Group A, B, and C). Adjudication of Test 3 was performed by the DRC sponsor member only.

Amendment of Planned Analyses: Exploratory evaluation of an optimal GH cut-off point for the MAC

Formally, Study 052 has not met one of two co-primary endpoints of the confirmatory efficacy analyses applying the predefined GH cut-off point for MAC and ITT. However, the results clearly indicate that optimal agreement of MAC and ITT outcomes would be obtained at a higher cut-off value, consistent with the observed higher potency of macimorelin for stimulating the GH release in this study. After a discussion of encouraging preliminary results from exploratory analysis suggesting superior performance of the MAC if evaluated at a higher GH cut-off point with FDA, it was considered acceptable by the Agency, that the sponsor would present in the CSR also results from exploratory analyses supporting the validity of the MAC at a recommended GH cut-off point and blood sampling scheme derived from Study 052.

Criteria to be considered when defining an optimal GH cut-off point for the MAC included the following

- Percent negative agreement of MAC with ITT lower CI limit above 75%.
- Percent positive agreement of MAC with ITT lower CI limit above 70%
- High repeatability of the MAC in the core and in the repeatability study
- High Sensitivity and specificity of the MAC
- Sensitivity analyses excluding data from subjects in the mITT population which are very likely invalid (i.e., data from only one subject (RS01-06) with apparent non-compliance or dosing error in the MAC of the core study)
- All above criteria should also be subjected to the planned hierarchical testing evaluating the different 'sparse sampling' scheme options

The performance characteristics of the MAC (negative, positive, and overall agreement with ITT, and repeatability of MAC were calculated for selected GH cut-off points including the pre-defined value of 2.8 ng/mL and evenly spaced values in a 'range of potential interest' ranging from 4.6 to 8.1 ng/mL, i.e., 2.80, 4.60, 5.10, 5.60, 6.10, 6.60, 7.10, 7.60, and 8.10 ng/mL, that includes 5.10 ng/mL as the pre-defined cut-off point for the ITT and 7.10 ng/mL as a value reflecting the 1.4-fold higher mean peak GH levels in MACs compared with ITTs. Performance characteristics were calculated based on peak GH concentrations based on all blood samples taken and for the predefined sampling windows for the planned hierarchical testing.

Analytical methods and analytical laboratories:

a) Macimorelin plasma concentrations were measured with a validated liquid chromatography-mass spectrometry (LC-MS/MS, detection limit of 0.2 ng/mL) at Prolytic GmbH, Germany.

b) GH serum concentrations were measured with a validated immunochemiluminometric assay (IDS-iSYS Human Growth Hormone, Immunodiagnostic Systems Ltd., UK), standardized to the recombinant GH calibration standard WHO 98/574, according to recommendations on assay standardization. Analytical laboratories: Europe: Synevo Central Lab Sp. z o.o., Poland; USA: LabConnect, LLC, TN, USA.

Number of subjects (see table below)**:**

Planned: At least 55 'ITT-positive' plus 55 'ITT-negative' subjects having completed the cross-over. Since it was unlikely that the number of ITT-positive and of ITT-negative subjects would be equal among the first 110 subjects with valid tests outcomes, the actual number of subjects needed to complete the study was likely to be greater than 110.

With regard to likelihood of AGHD, the following distribution was targeted:

Group A: High likelihood of AGHD (approx. 25% of the study population)

Group B: Intermediate likelihood of AGHD

Group C: Low likelihood of AGHD (approx. 25% of the study population)

Group D: Healthy controls (20-25 subjects matching a Group A subject)

At least 25% of the study patients were planned to be recruited at US sites.

Enrolled, SAF and mITT populations: 166 subjects were enrolled. 157 subjects received at least one dose of study drug and formed the safety population (SAF). Of 157 SAF subjects, 17 subjects did not fulfill the mITT criterion, i.e., randomized subjects in whom both GHSTs of the cross-over were evaluable. 140 subjects formed the mITT population; 31 (27%) of the 115 suspected AGHD patients (Group A, B, C) were treated in the USA. All healthy control subjects (Group D) were treated in a Phase I Unit in Poland. Number of patients was well distributed among AGHD likelihood Groups A, B and C with 42/38, 42/37 and 44/40 of the SAF/mITT population.

Parameter	N Planned	N Enrolled	N SAF	N mITT
All	at least 110	166	157	140
ITT-positive	at least 55	n.a.	n.a.	74
ITT-negative	at least 55	n.a.	n.a.	66
Group A (High likelihood of AGHD)	25%(26-28)	47	42	38
Group B (Intermediate likelihood of AGHD)		45	42	37
Group C (Low likelihood of AGHD)	25% (26-28)	45	44	40
Group D (Healthy controls)	20-25	29	29	25
N = number, n.a. = not applicable				

Main criteria for inclusion:

1. Male or female, aged between 18 and 65 years
2. Suspected growth hormone deficiency (GHD), based on either of the following:
   - Structural hypothalamic or pituitary disease, or
   - Surgery or irradiation in these areas, or
   - Head trauma as an adult, or
   - Evidence of other pituitary hormone deficiencies, or
   - Idiopathic childhood onset GHD (without known hypothalamic or pituitary lesion or injury).

OR (recruitment at a dedicated Phase I Unit only)

3. Group D: Healthy control

Subject matching a Group A subject by sex, age (± 5 years), BMI (± 2 kg/m^2), and estrogen status (women only).

Test product, dose and mode of administration, batch number:

Macimorelin (AEZS-130) oral solution 0.5 mg/mL for oral intake of 1 mL (0.5 mg) per 1 kg body weight. Batch number: 4001V (US); 4001V_B (8 European countries). Only one batch was used.

Duration of treatment:

Single dose macimorelin on day of the macimorelin GHST ('MAC'). One or two injections of regular human insulin on day of ITT (second injection in case of hypoglycemia not reached).

Reference therapy, dose and mode of administration, batch number:

Regular human insulin was obtained from pharmacy stock. In Poland and Serbia, the local provider distributed insulin. Intravenous single dose injection of 0.10 U/kg, lower and higher doses were also allowed (according to clarifying Note to File of 05 Apr 2016); recommended dose in subjects with a BMI > 30 kg/m^2 was 0.15 U/kg.

Injection of an additional insulin bolus of 0.05 U/kg if the target glucose value of less than (<) 2.2 mmol/L (40 mg/dL) AND symptomatic hypoglycemia were not achieved within 45 minutes after the initial insulin dose.

Batch numbers: Recorded in Patient File and insulin accountability logs.

Criteria for evaluation:

Efficacy

Primary:

Co-primary efficacy variables were 'Percent Negative Agreement' and 'Percent Positive Agreement' when using predefined cut-off points of both GHSTs.

Secondary:

Secondary efficacy criteria were 'Percent Overall Agreement' and estimated sensitivity and specificity of both GHSTs (macimorelin and ITT) when using predefined cut-off points of both GHSTs.

Safety:

Adverse events, vital signs, physical examinations, clinical laboratory investigations, and 12-lead ECG.

Other criteria were:

Test acceptance/preference by study subjects and investigators.

Preliminary PK: t_{max} and C_{max} of macimorelin plasma concentrations in the sampling period.

Preliminary PK/PD: t_{max} for macimorelin versus t_{max} for GH; C_{max} for macimorelin versus C_{max} for GH.

Amendment 1 (repeatability extension), efficacy variables:

Primary efficacy variable: Comparison of peak GH levels following repeated treatments with macimorelin.

Key secondary variable: Comparison of the outcome for both macimorelin treatments based on the pre-defined cut-off point, i.e., percent of positive, negative, and overall agreement.

Statistical methods:

All parameters have been presented by descriptive statistics with N, mean, SD, min, lower quartile, median, upper quartile and max for continuous data. Categorical data have been presented showing absolute frequencies and percentages.

Based on the mITT population the primary efficacy and the key secondary efficacy measures (percent negative and percent positive agreement) were analyzed confirmatory by a hierarchical testing procedure with regard to the sampling time for the macimorelin test:

1. Peak GH among all post baseline samples;
2. Highest GH among 60 and 45 minutes post dose;
3. GH at 60 minutes post dose;
4. GH at 45 minutes post dose.

The performance of the GHST with macimorelin would have been accepted as sufficient if the lower bound of the two-sided 95% confidence interval for the primary efficacy variable was 75% or higher for 'percent negative agreement' and 70% or higher for 'percent positive agreement'.

To control for overall Type I error rate, sufficient agreements for a method could have been claimed only when sufficient agreements can be claimed for all prior methods, if any.

ROC curves for sensitivity and specificity (and two-sided 95% CI) were presented assuming all high likelihood AGHD subjects as 'true' AGHD subjects and all healthy matching subjects as 'true' AGHD negative subjects.

After the confirmatory analysis based on the pre-defined cut-off point for the MAC had failed to meet the acceptance criterion for percent positive agreement, the second of the co-primary endpoint, while multiple secondary endpoints indicated clinically relevant advantages of the MAC over the ITT, exploratory analyses were discussed and agreed upon with FDA to characterize the performance characteristic of MAC when using different cut-off points. Also, the goal was to propose an "optimal" cut-off point to be recommended for clinical use.

Results - Conclusions:

Baseline characteristics:

Of 157 SAF subjects, 59% were male, 41% female, and 86% of white and 3%, 2%, 1% and 8% of Asian, Black/African American, Pacific Island or other origin. At screening, the median parameters were for age 41 years (range: 18 – 66 years), height 170 cm (range: 140 – 195 cm), weight 82 kg (range: 44 – 123 kg) and body mass index 28 kg/m^2 (range: 16 – 40 kg/m^2).

Out of these 157 subjects of the SAF, 17 subjects did not fulfill the criterion for being included in the modified intention-to-treat (mITT) population, i.e., not both GHSTs of the cross-over were evaluable. Thus, 140 subjects formed the mITT: 38 (27.1%) Group A, 37 (26.4%) Group B, 40 (28.6%) Group C, and 25 (17.9%) Group D.

Among the 115 suspected AGHD subjects in Groups A – C, 31 (27.0%) were included in the US.

Following their completion of the core study procedures, 34 patients from the mITT were included in the repeatability extension (Amendment no. 1) and had their MAC repeated ('Test 3').

Efficacy results:

Primary endpoint:

Based on the assessments (positive/negative) for MAC and ITT, the negative agreement was 93.94% and the positive agreement was 74.32%. The lower limit of the 95% CIs for the negative agreement was 85.20% and thus conformed with the preset criterion of 75% for this parameter.

Positive agreement of MAC with ITT was 74.32%, with a lower limit of 62.84% for the 95% CIs, which did not match the preset criterion of ≥ 70 % for this parameter.

Therefore, one of the two co-primary endpoints did not show sufficient agreement and the target endpoint of this study based on the pre-defined cut-off point for macimorelin was not achieved.

Growth Hormone Serum Concentrations:

Overall, mean/median peak GH levels correlated well with the likelihood of having AGHD as the highest GH values were determined in healthy subjects Group D (median peak GH = 16.1 ng/mL) and patients of Group C (median peak GH = 14.5 ng/mL), and the lowest GH levels were analyzed in Group A (median peak GH = 0.1 ng/mL).. Macimorelin induced approximately 1.4-fold higher GH concentrations than were obtained with the ITT. When compared on subject level, in about 80% of all cases peak GH levels following administration of Macimorelin were equal or higher than observed during the ITT.

Macimorelin Plasma Concentrations:

The mean maximal observed macimorelin plasma concentration (C_{max}) of 10.63 ng/mL was reached at mean t_{max} = 48.5 minutes (n=138); median C_{max} was 9.38 ng/mL and median t_{max} was 45.0 minutes. Highest values for mean/median macimorelin plasma concentrations per scheduled time point were determined at 45 and 60 minutes post dose with 8.8 ng/mL/7.4 ng/mL and 8.5 ng/mL/7.1 ng/mL. These are the same time points when highest total GH concentrations were measured. No significant differences in the mean C_{max} for macimorelin per AGHD likelihood group was observed.

Sensitivity and Specificity Evaluation in Group A and Group D Subjects:

Sensitivity (Figure 1) and specificity (Figure 2) for both GHSTs were estimated, assuming all high likelihood AGHD subjects of Group A as 'true' AGHD subjects and all healthy matching subjects of Group D as 'true' AGHD negative subjects. When using the pre-defined cut-off points of 2.8 ng/mL for the MAC

and 5.1 ng/mL for the ITT, point estimates for sensitivity ranged from 0.87 to 0.90 for the MAC and from 0.97 to 1.0 for the ITT, depending on the inclusion or exclusion of data from not matched Group A subjects, respectively. For both GHSTs, the estimated specificity was 0.96, irrespective of the in/exclusion data from not matched Group A subjects. Based on such data incl. related ROC (receiver operator characteristics (ROC) curves, it can be concluded that the pre-defined cut-off point of 5.1 ng/mL for the ITT seemed to be appropriate resulting in very good sensitivity and specificity.

Figures 1 and 2 illustrate the effect of varying the GH cut-off point for the MAC on the estimated sensitivity and specificity, respectively. Figure 1 shows that increasing the GH cut-off point for the MAC between 2.8 ng/mL and about 8 ng/mL will increase in the sensitivity with no or only a minimal decrease in the specificity, as evident from the flat profile in Figure 2.

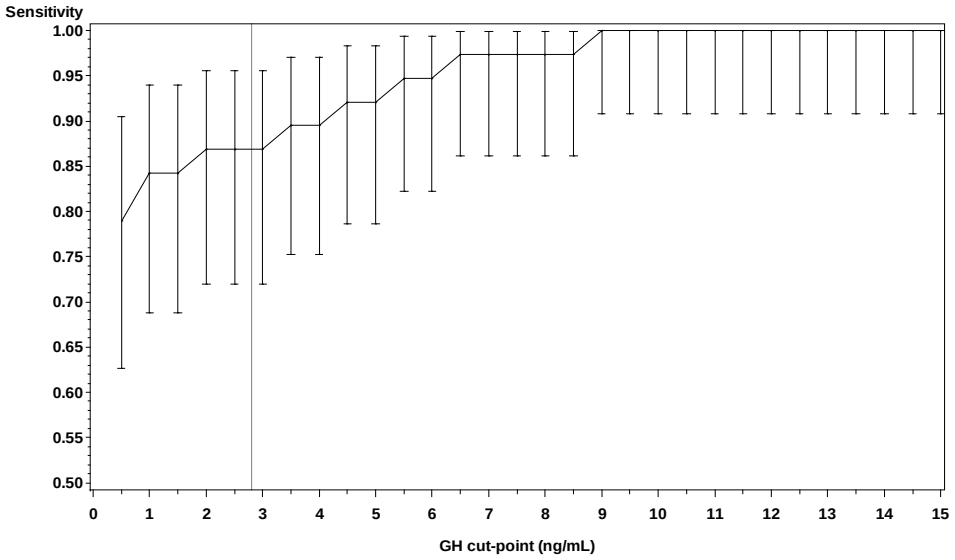

Figure 1: MAC: Sensitivity for varying GH cut-off points of group A and D subjects

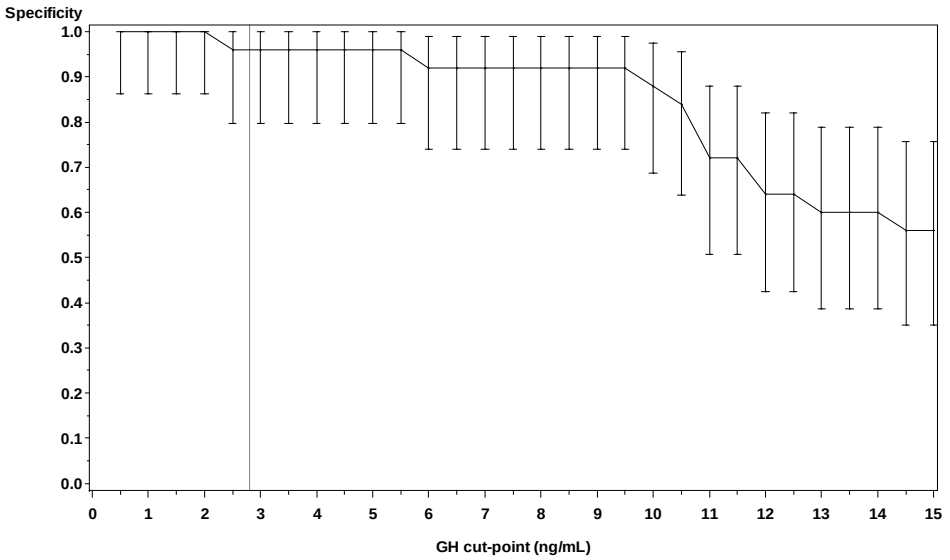

Figure 2: MAC: Specificity for varying GH cut-off points of Group A and D subjects

Questionnaire Study Results:
The vast majority of subjects, 95.5% if asked after the MAC and 90.4% if asked after the ITT as their second GHST, would choose the MAC if they would need to undergo another GHST in the future.

Repeatability of MAC:
Growth hormone serum concentrations of the two MACs performed for repeatability matched well, both in terms of peak levels and time course of the GH values. In 32 out of 34 (94%) planned repetitions of a MAC, the test outcome was repeatable at the pre-defined cut-off point of 2.8 ng/mL.

Exploratory Analyses for the GH cut-off point:
Exploratory analyses showed that the study data could support the recommendation of a range of cut-off points for MAC, up to an upper limit of 8.1 ng/mL (see Table 1).

Table 1: Agreement, repeatability, and sensitivity and specificity for GH cut-off points 'of interest' preselected based on exploratory analyses based on peak GH values in MAC excluding data from RS01-06 (compliance issue)

GH cut-off point	**Core study analysis** **Agreement between MAC and ITT** (mITT ***w.o.RS01-06; N=139)***						**MAC Repeatability** (M-core vs M-rep.; ***w.o.RS01-06; N=33)***		**ROC analysis for MAC** (Groups A+D; N=38+25)	
	Negative agreement		**Positive agreement**		**Overall agreement**		**Overall agreement**		**Sensi-tivity**	**Speci-ficity**
	(%)	Lower CI limit (%)	(%)	Lower CI limit (%)	(%)	Lower CI limit (%)	(%)	Lower CI limit (%)	(%)	(%)
2.80	95.38	**87.10**	74.32	62.84	84.17	77.02	**96.97**	**84.24**	87	**96**
4.60	95.38	**87.10**	81.08	**70.30**	87.77	**81.14**	**93.94**	**79.77**	**92**	**96**
5.10	93.85	**84.99**	82.43	**71.83**	87.77	**81.14**	**93.94**	**79.77**	**92**	**96**
5.60	93.85	**84.99**	85.14	**74.96**	89.21	**82.83**	**93.94**	**79.77**	**95**	**96**
6.10	92.31	**82.95**	86.49	**76.55**	89.21	**82.83**	87.88	71.80	**95**	**92**
6.60	89.23	**79.06**	87.84	**78.16**	88.49	**81.98**	**90.91**	**75.67**	**97**	**92**
7.10	89.23	**79.06**	89.19	**79.80**	89.21	**82.83**	**90.91**	**75.67**	**97**	**92**
7.60	89.23	**79.06**	89.19	**79.80**	89.21	**82.83**	**90.91**	**75.67**	**97**	**92**
8.10	87.69	**77.18**	91.89	**83.18**	89.93	**83.68**	**90.91**	**75.67**	**97**	**92**

Additional analyses, including clinical considerations, led to the conclusion that 5.1 ng/mL is an appropriate GH cut-off point to be recommended, which limits the risks of overdiagnosis.

Table 2 summarizes the main performance characteristics of the MAC for this GH cut-off point of 5.1 ng/mL.

Table 2: Agreement, repeatability, and sensitivity and specificity explored for MAC GH cut-off point of 5.1 ng/mL in hierarchical testing with or without excluding data from RS01-06 (compliance issue)

Hierar-chical testing **Step** (w.o.: RS01-06 data excluded)	**Core study analysis Agreement between MAC and ITT** (mITT=140) (w.o. RS01-06: N=139)				**MAC Repeatability (M-core vs M-repeat.)** (N=34) (w.o. RS01-06: N=33)		**ROC analysis for MAC (Groups A+D)** (N=38+25))	
	Negative agreement		**Positive agreement**		**Overall agreement**		**Sensi-tivity**	**Speci-ficity**
	(%)	Lower CI limit (%)	(%)	Lower CI limit (%)	(%)	Lower CI limit (%)	%	%
1	92.42	**83.20**	82.43	**71.83**	**91.18**	76.32	92	96
2	89.39	**79.36**	83.78	**73.39**	**91.18**	76.32	**92**	**96**
3	89.39	**79.36**	87.84	**78.16**	88.24	72.55	**95**	**96**
4	80.30	68.68	85.14	**74.96**	88.24	72.55	**95**	**88**
1 w.o.	93.85	**84.99**	82.43	**71.83**	**93.94**	79.77	92	96
2 w.o.	90.77	**80.98**	83.78	**73.39**	**93.94**	79.77	92	96
3 w.o.	90.77	**80.98**	87.84	**78.16**	**90.91**	75.67	95	96
4 w.o.	81.54	69.97	85.14	**74.96**	**90.91**	75.67	95	88

Subgroup Analyses:

From all 140 subjects of the mITT population, 27 (19%) and 15 (11%) subjects had a BMI of 30.0 – 34.9 kg/m² and 35.0 – 40.0 kg/m². No healthy control subject was enrolled with a BMI ranging 35.0 – 40.0 kg/m². 24 (17%) subjects were in the age range of 18 to ≤ 25 years. The number of subjects included into the analyses of these subgroups was too small to obtain meaningful statistical results.

83 subjects of the ITT were male, 57 female. Within the same subgroup of Group A and Group D subjects, the analyses showed higher sensitivity of the MAC at the recommended cut-off point of 5.1 ng/mL both for male and female subjects (0.91 and 0.94) than at the predefined cut-off point (0.86 and 0.88). The values for specificity of 1.0 in male and 0.91 in females were not different between both cut-off points

Safety results:

Any test emergent adverse event (TEAE) was recorded in 25.3% of a total of 154 subjects following MAC (combined core study and repeatability extension), in contrast to 96.2% of a total of 157 subjects following ITT, with a total number of 77 and 761 TEAEs, respectively.

One **serious** TEAE **(SAE)** was reported in this study after the first MAC (upper limb fracture due to fall from a ladder) and which was reported unrelated to study drug.

One **severe** TEAE was recorded in 1 subject following MAC, in contrast to 25 severe TEAEs in 11 (7%) of a total of 157 subjects following ITT. Severe TEAEs after the ITT included: 5 (3.2%) subjects each with somnolence, hyperhidrosis; 4 (2.5%) with asthenia; 3 (1.9%) with hunger; 2 (1.3%) with nervousness, and 1 (0.6%) with tremor.

Moderate TEAEs that were reported during/after MAC for 2 of 154 (1.3%) patients each included the following: nasopharyngitis, upper respiratory tract infection, headache, and fatigue. A total of 229 moderate TEAEs during/after the ITT were reported for 64 (40.8%) subjects.

Altogether, 57 mild TEAEs were recorded in 25 (16.2%) of 154 subjects following MAC and 507 mild TEAEs in 76 (48.4%) of a total of 157 subjects following ITT,

Any TEAE with likely or possible causal relationship had been reported for 14.3% of the 154 subjects following a MAC, in contrast to 94.9% of 157 subjects following an ITT.

For MAC, dysgeusia (4.5% of subjects) was the most frequently reported TEAE likely or possibly related to macimorelin, followed by fatigue ((3.2%), headache and nausea (2.6% each).

No medically significant abnormality was recorded in any parameter of clinical laboratory investigations, vital signs or physical examination. No test discontinuation or premature termination of the study was noted for the MAC.

No significant difference in the spectrum of TEAEs reported for the MAC core study as well as for MAC repeatability was noted.

Overall, macimorelin was safe and well tolerated. The frequency and severity of observed TEAEs show that the MAC was associated with minimal side effects when compared with the ITT that was associated with a broad spectrum of signs and symptoms including also moderate and severe events related to hypoglycemia.

Conclusions:

In the planned analyses, the study reached the predefined statistical acceptance criterion for 'percent negative agreement', the clinically more important one of the co-primary efficacy variables. However, the study failed to reach the acceptance criterion for the 'percent positive agreement' at the predefined value of 2.8 ng/mL for the GH cut-off point for the MAC.

However, the study results clearly support the conclusion that compared with the ITT, the MAC has distinct advantages in several clinically relevant features:

- **Feasibility and Robustness**: MAC does not rely on procedures and criteria that are difficult to achieve and that might require a repetition of a MAC. In 17% of 157 ITT subjects, hypoglycemia has not been achieved and required test repetition, and 4 tests (24%) from 17 repeated ITTs did not provide an evaluable result even on 'second try'.
- **Repeatability**: MAC outcome was highly repeatable, both in terms of stimulated GH concentrations and classifications of the test outcome.
- **Safety**: MAC was associated with a very low rate of likely or possibly drug-related adverse events (affecting 14.3% of the test subjects) and that were mostly mild. As a prerequisite for evaluability, ITT requires the achievement of hypoglycemia that is - intentionally - associated with a high rate of symptoms (94.9%), including potentially severe symptoms (affecting 7% of the test subjects), and therefore requires close safety monitoring by medical staff.
- **Acceptability**: 95.5% of the subjects who were asked after the MAC and 90.4% who were asked after the ITT as their second GHST declared MAC as first choice for future assay.
- **GH-provocative potential**: A disagreement of MAC and ITT outcome was in most cases related to consistently – on average 1.4-fold - higher peak GH concentrations in MAC.

Supplementary exploratory analyses showed that the original study data could support the recommendation of a range of cut-off points for MAC, up to an upper limit of 8.1 ng/mL. Conclusions and recommendations below are based on sensitivity analyses excluding data from an obviously non-compliant or incorrectly dosed subject (RS01-06).

Based on the following considerations, a value of 5.1 ng/mL is finally concluded to be the recommended optimal GH cut-off point for the MAC, with collection of blood samples for GH measurements 45 and 60 minutes after intake of the macimorelin test dose:

- **Agreement of MAC and ITT outcome**: Both negative and positive agreement are clearly within the acceptance limits for this cut-off point, with point estimates of 90.8% and 83.8% and lower limit of the 95% confidence interval of 81% and 73.4%, respectively.
- **Repeatability:** High repeatability of MAC, with a point estimate of 93.9%
- **Sensitivity and specificity**: High sensitivity (92%) and specificity (96%).
- **Sampling window**: Hierarchical testing indicates that the above favorable performance characteristics of MAC are achieved when the evaluation is based only on the stimulated GH levels measured 45 and 60 minutes after the test dose and are maintained even for an evaluation on the single value measured 60 minutes after the test dose only. In the interest of a most reliable test outcome, however, a performance of the MAC with 2-point blood sampling at 45+60 minutes post-dose may be more appropriate. Additional GH measurements post-dose or at other time-points post dose are not required to assure evaluability or outcome of MAC.

• **Clinical considerations**: The recommended cut-off point for MAC limits the risk of over-diagnosing AGHD. In summary, MAC is a feasible, robust, safe, and highly reproducible test to diagnose AGHD, being a better means for evaluating a patient with suspected AGHD than the ITT
Date of report: 16-June-2017