STATISTICAL REPORT

Clinical trial evaluating the effectiveness and safety of Dvectis Single and Dvectis Double pads in comparison to “no pad use” in patients with chronic lumbar spine pain

Protocol code: DVE-17

Tested devices	Dvectis Single Dvectis Double
Indication	Chronic lumbar spine pain
Design	Randomized controlled clinical trial
Study start	9 Apr 2018
Study end	20 Aug 2018
Principal investigator	MUDr. Martin Holinka Department of Orthopaedics Karvinská hornická nemocnice a.s. Zakladatelská 975/22 735 06 Karviná - Nové Město
Contracting entity	DD - Hippero s.r.o. třída Kpt. Jaroše 1922/3 602 00 Brno
Document author	PharmTest s.r.o. Hradec Králové

DOCUMENT VERSION

Document version: Final
Version date: 07-12-2018

DOCUMENT VERSION HISTORY

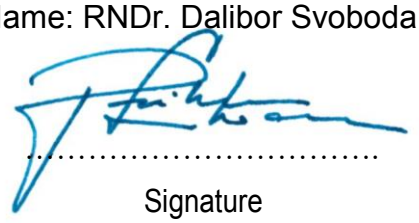
Version	Version date	Author	Commentary
Final	07-12-2018	Dalibor Svoboda	Formal finalization only

SIGNATURE PAGE

STUDY STATISTICIAN

Name: RNDr. Dalibor Svoboda

PharmTest s.r.o.



7. 12. 2018

…….
Signature

…….
Date

CONTENTS

1 List of Abbreviations

ANOVA	Variation analysis
ARS	Analysed population of all randomized study subjects
CRF	Case Report Form
H0	Null hypothesis
HA	Alternative hypothesis
RI	Reliability interval
ITT	Analysed population defined for maximum observance of the *Intention*-To-Treat principle, i.e. including all subjects with serious deviations from the study protocol
LOCF	Last Observation Carried Forward (data imputation method)
NOCB	Next Observation Carried Backward (data imputation method)
MCAR	*Missing Completely At Random* (type of missing data distribution)
PI	The most intensive pain felt in the last 48 hours and recorded based on VAS
PID	Pain Intensity Difference *(PI* difference compared to week 1)
PPS	Analysed population of study subjects without serious deviation from the study protocol
SAF	Population analysed for safety assessment
SAS	Software products of SAS Institute
SD	Standard deviation
UMVUE	Uniform Minimum Variance Unbiased Estimator
VAS	Visual analogue scale
WCS	Worst Case Scenario (data imputation method)
Z	*Z*-statistics

2 Introduction

2.1 Document Purpose and Objective

Presentation of results of the clinical trial data statistical analysis.

2.2 Description of the Trial, Objectives, Hypotheses

2.2.1 Trial Description

The trial was a monocentric, open, randomized, 3-arm clinical trial that took place at the Department of Orthopaedics of the Karviná Miners' Hospital (Karvinská hornická nemocnice a.s.). The involvement of a maximum total number of 198 patients with chronic lumbar spine pain was planned; the patients were randomly and evenly assigned to one of the 3 treatment groups:

1. Use of the Dvectis Single pad;
2. Use of the Dvectis Double pad;
3. Use of no pad.

One sequential interim analysis was planned in the middle of the clinical trial and after this interim analysis, the trial was stopped early due to proven efficacy.

The expected period of participation of each patient in the clinical trial was 6 weeks (± 5 days). For an overview of the individual study visits and procedures, study population and other detailed information, see the study protocol.

According to the plan, the assessment subject was asked during Visit 1 and Visit 2 about the most intensive pain felt in the last 48 hours (PI). The pain was recorded by the subject in the visual analogue scale (VAS) in CRF under supervision of the investigator. Then the assessment subject recorded the intensity of their pain independently in the Patient Journal daily.

2.2.2 Primary Objective, Quantity and Hypothesis

The primary objective of the clinical trial was to assess the efficacy of using the Dvectis Single pad in comparison to “use of no pad” in patients suffering from chronic lumbar spine pain.

PI was the primary quantity.

The primary hypothesis was the superiority of Dvectis Single based on a check, assessed based on the difference in *PI* (*PID*) between week 2 and week 6 (*PID6*).

2.3 Detected Changes in Comparison to Statistical Analysis Planned in Study Protocol

Contrary to the study plan, the clinical trial took place in a single study centre. Contrary to the plans included in the study protocol, the factor of the study centre was not accounted for in the analyses in any way

and the presence of interaction between treatment and the study centre was not analysed.

For the same reason, the plan to possibly combine small centres was not used either.

There were no other changes detected in the plan of the first sequential interim analysis.

2.4 Detected Changes in Comparison to Statistical Plan

The statistical plan was prepared before the statistical analysis performance and formally approved. It included technical details for the statistical analysis performance and the method of result presentation.

The only deviation from the statistical plan is related to the summaries for the treatment regimen compliance assessment (see Table 3 Treatment regimen compliance). The original plan was based on the evaluation of cumulative sitting time. However, the average times were analysed in reality. This is due to the fact that the cumulative (or individual) sitting times were not included in the data sheet, therefore they were not included in the analytical datasets either. From this point of view, it is an error in the statistical analytical plan without a practical effect.

3 Analysed Populations and Number of Subjects

Prior to the statistical analysis, 4 populations were defined in line with the study protocol:

1. Population including all randomised subjects (ARS);
2. Population of all patient with intention to treat (ITT);
3. Population of patients without serious deviation from the clinical trial plan (Per-Protocol Set
   – PPS);
4. Safety population.

Each analysed population was based on a list of its members. Although the treatment group could have generally been different for each patient in a different population (see the study protocol and statistical analysis plan), such a case did not occur (as all the patients were *de facto* placed in the group to which they were randomized). Therefore the real treatment group is the same as the randomization treatment group for all patients. This fact is not discussed further in the document.

3.1 ARS Population

The ARS population included all patients who were involved in the clinical trial and randomized. The ARS population was used for summaries of demographic quantities and other quantities established during Visit 1 in order to describe the study population.

3.2 ITT Population

In this case, this population was identical to the ARS population. According to the plan, ITT was not supposed to include those ARS patients for whom no *PI* data were available and those patients who entered the study repeatedly. However, such circumstances were not detected. The ITT population represented the primary analysed population for the assessment of efficacy.

3.3 PPS Population

According to the statistical plan, PPS was to include all ITT patients except for patients with a serious deviation from the clinical trial plan; a serious deviation includes:

1. Failure to meet one or multiple participation criteria;
2. Error in treatment allocation;
3. Documented use of unauthorized concomitant medication or treatment;
4. Missing *PI0* value;
5. More than 25% missing *PI* values.

6. Serious treatment regimen violation or another serious deviation from the clinical trial plan.

The monitoring reports and discrepancy report implied:

1. No study subject that has not met a participation criterion has been registered;
2. No study subject with incorrectly allocated treatment has been identified;
3. No use of unauthorized concomitant treatment has been documented;
4. There was no *PI0* value missing;
5. Patients 45, 53 and 94 ended the study early; the number of *PI* values missing is over 25% (83%).
6. The following patients violated the treatment regimen in terms of non-observance of the minimum time of sitting on the pad (21, 23, 35, 42, 55, 57, 60, 72, 75, 76, 84, 86, 88).

Thus there was a serious deviation from the protocol in 16 study subject, i.e. 15.7%.

The PPS population represented the secondary analyzed population for the assessment of efficacy. Its task was to demonstrate the robustness of the main clinical trial results. For this reason, comparisons of the tested group efficacy to the control group were used (in addition to ITT) in final analyses using *PI* in order to achieve confirmation of the main clinical trial conclusions.

3.4 Safety Population

According to the plan, the safety population should have constituted of subjects from ARS, except for those with whom all contact was lost immediately after Visit 1 and those who were placed (in reality) in one of the tested groups but were not given a pad.

A pad was given to all patients in the tested groups. All contact was lost with patients 53 and 94 and they were eliminated from the safety population.

Patients in the safety population were analyzed based on the treatment groups they were really placed in, regardless of what group they had been randomized to.

3.5 Method of Defining Analyzed Populations for Final Analysis

The required analyzed populations were defined before the analysis performance according to the plan.

3.6 Overview of Analysed Populations and Patient History

Table 1 Numbers of subjects in treatment groups and analysed populations

	Dvectis Single	Dvectis Double	No treatment	Total
ARS population	35	34	33	102
ITT population	35	34	33	102
PPS population	29	25	32	86
Safety population	35	33	32	100

3.7 Selection Size Calculation

The selection size calculation was based on the estimated difference in efficacy expressed by the primary quantity between the Dvectis Single group and the control group, which was 15 mm. Due to the degree of uncertainty in expectations and considering the fact that even differences of approximately 10 mm are of obvious practical and clinical significance, the interim analysis was planned to allow the detection of this minor difference strongly enough, even if the original expectations had been too optimistic.

The following assumptions were applied to the determination of the number of patients:

1. The primary hypothesis is based on a comparison of to arms, namely the Dvectis Single pad and the control group. For this purpose, the Dvectis Double arm plays no direct role in the selection size.
2. The null hypothesis is based on zero difference between the compared groups.
3. The alternative hypothesis corresponds to the difference of 10 mm in favour of the Dvectis Single pad.
4. The spread of results is the same in both compared groups, is not efficacy-dependent and corresponds to the standard deviation of 15 mm. This value was estimated based on previous studies.
5. The null hypothesis test is unilateral, at the significance level of $\alpha = 0.025$.
6. Test strength of 90% is required to detect the difference corresponding to the alternative hypothesis (regardless of the stage when the clinical trial will be completed).
7. At most 2 sequential analyses will be performed, equidistantly. The first will be performed in the middle of the expected maximum number of study subjects.
8. After the first sequential analysis (interim analysis), the study may be ended due to efficacy (success) based on the clearly defined completion criteria stated above.
9. The test statistics limits have been determined according to Pocock (i.e. they are the same for both sequential analyses and guarantee the maintenance of cumulative probability of type I error occurrence at the previously defined level of significance specified in (5).

Implementation of the above assumptions resulted in the requirement of the maximum of 52.02 assessable subjects in each arm. To comply with the requirement of the clinical trial balance (i.e. the same number of subjects in the Dvectis Double pad group), a total of 3×52.02=156.06 of assessable subjects was required.

Although the primary analysis was performed on the analyzed ITT group, due to the requirement of the clinical trial robustness demonstration in terms of conclusion insensitivity to the analyzed population selection and deviations from the clinical trial plan, the concept of assessability needs to be related to the analyzed population without any serious deviation.

Taking the assumption of no more than 20% of subjects with a serious deviation from the clinical trial plan into account, the total of approximately 156.06/0.8=195.08 subjects had to be included in the clinical trial. The closest higher integer divisible by 6 (three groups x 2 analyses) is 198. Therefore the maximum of 66 subjects in each treatment group had to be included in the clinical trial and the interim analysis had to be performed at the moment when 33 subjects were completed in the smaller of the two groups – the Dvectis Single group and the control group.

4 Missing Data

In the ARS population analyses, the missing data were be treated as missing according to the plan, i.e. stated in summaries as total numbers of missing data; however, they were excluded from the calculations of other statistics. In reality, only the time of sitting on the pad is of significance.

In all analyses of the *PI* quantities, the method of imputation of the last observed value (*Last* Observation Carried Forward – LOCF) was used, in accordance with the plan. The plans for a situation when the *PI1* would be missing were not used since the situation did not occur.

The use of LOCF generally also leads to a bias but with regard to the observed decrease in the PI values during the patient's participation in the study in the tested groups, compared to the reference group, such an approach is obviously conservative.

For evaluation of the Oswestry questionnaire, the use of LOCF was planned primarily at the level of individual questions as well as the use of the scoring manual procedure accounting for missing values, if any. This procedure was beyond the data administration and statistical analysis activities since the questionnaire was evaluated by the investigator, entering only the total score in the patient record sheets.

If the total score was missing, it was processed similarly to the processing of missing values in the ARS population analysis (i.e. solely elimination from the analysis).

5 Sequential and Final Analysis

The first sequential analysis was performed according to the interim analysis plan and resulted in early study stopping due to efficacy. A report of the first sequential analysis results is included in a separate document. Then the final analysis was performed, which is presented in this report.

6 Characteristics of Study Subjects and Treatment Regimen Compliance

For the description of the study population within the sequential interim analysis, summaries of demographic quantities were calculated and presented, as established before the start of the treatment in the scope determined by the statistical analysis plan. The summaries were presented for the respective treatment groups (for all 3 treatment groups). The significance of the differences in these characteristics was not tested.

The treatment regimen compliance was evaluated by a calculation of aggregate statistics for the average time of the device use in the course of participation in the study.

6.1 Demographic Quantities and Other Quantities Measured before the Start of Treatment

6.1.1 Analysis

For the analyzed quantities, summary statistics were calculated in the respective treatment groups and as a total. The analyzed ARS population was used.

6.1.2 Presentation of Results

Table 2 Demographics and other quantities before the start of treatment

	Dvectis Single N = 35	Dvectis Double N = 34	No treatment N = 33	Total N = 102
Age (year)				
Minimum	29	26	37	26
Average (SD [1])	50.1 (9.1)	45.9 (9.9)	51.1 (7.5)	49.0 (9.1)
Median	52.0	47.0	52.0	50.0
Maximum	65	60	63	65
Total [2]	35	34	33	102
Sex, N (%)				
Male	3 (8.6%)	6 (17.6%)	7 (21.2%)	16 (15.7%)
Female	32 (91.4%)	28 (82.4%)	26 (78.8%)	86 (84.3%)
Total [2]	35	34	33	102
Body weight (kg)				
Minimum	43	52	53	43
Average (SD [1])	78.7 (18.3)	78.4 (13.7)	78.3 (16.8)	78.5 (16.3)
Median	75.0	78.5	78.0	77.5
Maximum	120	103	123	123
Total [2]	35	34	33	102
Body height (cm)				
Minimum	153	148	158	148
Average (SD [1])	168.1 (8.3)	168.4 (8.0)	168.7 (7.3)	168.4 (7.8)
Median	168.0	168.0	168.0	168.0
Maximum	189	185	187	189
Total [2]	35	34	33	102
Pain intensity before treatment (mm VAS)				
Minimum	45	45	43	43
Average (SD [1])	62.6 (14.2)	63.0 (13.7)	59.5 (12.8)	61.7 (13.5)
Median	60.0	60.0	59.0	60.0
Maximum	98	90	96	98
Total [2]	35	34	33	102

Analyzed population: ARS Dataset: Baseline Data image: 2018-12-05

Comments:

1) SD = Standard deviation.
2) Number of analyzed data (without missing values).

6.2 Treatment Regimen Compliance

Table 3 Treatment regimen compliance

Dvectis Single N = 35	Dvectis Double N = 34	No treatment N = 33	Total N = 102
Average daily time of sitting on the pad			
Minimum	0.32	0.42	Not applicable	0.32
Average (SD [1])	3.496 (2.253)	2.864 (1.472)	Not applicable	3.185 (1.921)
Median	2.806	2.542	Not applicable	2.611
Maximum	8.69	6.28	Not applicable	8.69
Total [2]	34	33	0	67

Analyzed population: ARS Dataset: Diary Data image: 2018-12-05

Comments:

1) SD = Standard deviation.
2) Number of analyzed data (without missing values).

7 Efficacy

7.1 Primary Quantity

7.1.1 Basic Information

The primary quantity was pain intensity (*PI*) measured on a 100 mm visual analogue scale (VAS). Measurements were taken once a week in weeks 1-6 (*PI1, PI2, …, PI6*).
The basic aggregate statistics for each treatment group were calculated and presented individually each week.
The pain intensity difference (*PID*) was calculated from *PIi* in weeeks 2 – 6 based on the relationship:

$$PID_i = PI_1 - PI_i, \quad i = 2, ..., 6$$

The basic aggregate statistics for each treatment group were calculated and presented for the respective PID_i.
PID_6 was used for testing the primary hypothesis, statistically formulated in the clinical trial plan (protocol) as follows: If we mark μ_{DS} the mean value PID_6 for Dvectis Single and μ_{Ref} the mean value PID_6 for the control (reference) group, then the primary null hypothesis H0

$$H0: \quad \mu_{DS} - \mu_{Ref} \leq 0$$

is tested against the alternative HA

$$HA: \quad \mu_{DS} - \mu_{Ref} > 0$$

unilaterally at the significance level of α = 0.025.

The testing itself is not a part of the final analysis presented in this document since the test and the subsequent efficacy conclusion were made based on the sequential analysis results.

The analysis of *PID6* was performed in the ANOVA model. The model included the treatment factor and the *PI1* covariate. The results are presented in the form of a standard ANOVA table, including the sums of type III squares (in the SAS terminology).

7.1.2 Ordering

Ordering is significant for the performance of the final analysis. Testing of hypotheses, calculations of *p*-values and the construction of intervals were based on the notion that the resulting statistics can be ordered, the ordering criterion being their extremeness while the null hypothesis is valid. E.g. *the Z*-statistics is the more extreme the higher its (absolute) value. The performance of sequential analyses is complicated due to the fact the result is not a number but a pair of numbers (*M*, *Z*), constituted by the number of the interim analysis (stage) when the clinical trial is stopped within the performed sequential analyses and the calculated *Z*-statistics. Therefore a method of ordering the results by extremeness had to be defined. The ordering definition allows decisions to be made on the probability of inequality

for arbitrary *M1, M2, Z1, Z2.* However, the optimal method of ordering cannot be generally determined since the sequential statistics densities do not have a monotonic value relationship, therefore the selection criterion may not be applied with regard to the test with the highest power. In general, there are several intuitive options.

According to the plan in the study protocol and the statistical analysis plan, the results related to the primary hypothesis were ordered based on *stage-wise ordering (analysis time ordering / Fairbanks and Madsen ordering*) [1]. A number of authors consider this option intuitively attractive, for the following reasons, among others:

- *The P*-value obtained during the final analysis is lower than the level of significance α used in the sequential test in the interim analysis (after the first and second stage) only when the null hypothesis is rejected [2]. This feature is important in relation to the internal consistence of the results and is generally not demonstrated by other types of ordering.
- When the clinical trial is stopped early, the *P*-value is not based on the information quantity, or the number of patients that would be included if the trial was not stopped early [2].

7.1.3 Final Analysis Related to Primary Hypothesis (Dvectis Single Pad)

Clinical trial stopping based on the observation of an extreme result during sequential analyses diverts the point estimates made in a way common for a design with a fixed number of patients without sequential analyses. Detailed information is given in the study protocol and the statistics plan.

In this particular case, the clinical trial was stopped early after the **first** sequential interim analysis. This fact, with the application of the ordering described above, led to estimates identical to the naive 95% reliability interval calculated for the fixed selection size.

7.1.4 Final Analysis Related to Dvectis Double Pad Efficacy

The test of the Dvectis Double pad difference from placebo expressed an important secondary objective of the clinical trial, which was to allow additional statements concerning efficacy. The difference significance test could have been performed solely if the null hypothesis expressing the primary objective was rejected in any stage of the clinical trial. This condition was met after the first sequential analysis. The secondary hypothesis test was performed just once.

The condition of test performance solely if the primary null hypothesis is rejected is a type of hierarchy testing when it is easy to maintain the total cumulative probability of type I error occurrence (*familywise error rate*) at a previously specified level α simply by the fact both tests are performed at the same level α individually (hierarchically), which may be formally proven with the use of the *closed testing principle* [6]. However, this procedure cannot be applied to sequential interim analyses exactly, therefore in the case of this clinical trial, it does not allow strict control of error inflation probability α. Hung [7] proposes an intuitive strategy based on secondary hypothesis testing at significance level α, which corresponds to (1-*α*) percentile of the standard Gaussian distribution, if the primary null hypothesis was rejected at the **total** significance level α (i.e.

total within the sequential tests). Although it is obvious this strategy may not necessarily result in strict control of type I error probability inflation, it does not represent a serious problem in this case [8].

With regard to the above, the difference between the Dvectis Double pad and placebo was analyzed based on a unilateral superiority test at the significance level α = 0.025, using the calculation of a standard bilateral 95% reliability interval for the difference between the tested and the control group and the establishment of its position to 0. This test rejected the null hypothesis (see the Results part) and is conclusive since the primary null hypothesis was rejected before the test performance (for Dvectis Single).

7.1.5 Primary Quantity Exploratory Analysis

Within the exploratory analysis, the correlation between *PID6* and the recorded total average weekly time of sitting on the pad was investigated in the pad groups. Correlation coefficients were calculated and presented (Pearson correlation coefficient, Spearman order correlation coefficient ρ and Kendall's τ) together with *p*-values of the null hypothesis test values *correlation coefficient*=0. However, these *p*-values can serve only as indicators of interesting results worth further investigation; they are not aimed for drawing confirmatory conclusions.

7.1.6 Final Analysis of Other Clinical Trial Objectives

No other planned analyses were aimed for confirmatory purposes and they were processed regardless of the performed sequential analyses.

7.1.7 Presentation of Results

7.1.7.1 Aggregate Statistics for Primary Quantity and Related Derived Quantities

Table 4 PI summaries

	Dvectis Single N = 35	Dvectis Double N = 34	No treatment N = 33	Total N = 102
PI1 (mm)				
Minimum	45	45	43	43
Average (SD [1])	62.6 (14.2)	63.0 (13.7)	59.5 (12.8)	61.7 (13.5)
Median	60.0	60.0	59.0	60.0
Maximum	98	90	96	98
Total [2]	35	34	33	102
PI2 (mm)				
Minimum	6	17	1	1
Average (SD [1])	53.1 (23.9)	57.9 (13.5)	60.2 (19.4)	57.0 (19.5)
Median	60.0	59.0	59.0	59.0
Maximum	98	91	94	98
Total [2]	35	34	33	102
PI3 (mm)				
Minimum	2	26	4	2
Average (SD [1])	46.4 (23.5)	50.6 (12.8)	61.4 (17.7)	52.7 (19.4)
Median	52.0	50.5	60.0	55.0
Maximum	98	77	95	98
Total [2]	35	34	33	102
PI4 (mm)				
Minimum	0	14	3	0
Average (SD [1])	39.5 (24.6)	42.1 (14.7)	61.3 (20.1)	47.4 (22.3)
Median	45.0	42.0	61.0	49.5
Maximum	98	73	98	98
Total [2]	35	34	33	102
PI5 (mm)				
Minimum	0	3	4	0
Average (SD [1])	34.4 (22.3)	32.4 (17.3)	61.9 (20.6)	42.6 (24.1)
Median	30.0	28.5	65.0	42.5
Maximum	98	72	95	98
Total [2]	35	34	33	102
PI6 (mm)				
Minimum	2	1	2	1
Average (SD [1])	26.9 (23.5)	25.6 (16.9)	61.4 (18.2)	37.6 (25.6)
Median	20.0	19.5	62.0	39.0
Maximum	98	64	92	98
Total [2]	35	34	33	102

Analyzed population: ITT Dataset: VasPid Data image: 2018-12-05

Comments:

1) SD = Standard deviation.
2) Number of analyzed data (without missing values).

Table 5 PID summaries

	Dvectis Single N = 35	Dvectis Double N = 34	No treatment N = 33	Total N = 102
PID_2 (mm)				
Minimum	-29	-20	-32	-32
Average (SD [1])	9.5 (24.1)	5.1 (13.7)	-0.6 (19.7)	4.8 (19.9)
Median	2.0	0.0	-2.0	0.5
Maximum	73	40	84	84
Total [2]	35	34	33	102
PID_3 (mm)				
Minimum	-24	-10	-25	-25
Average (SD [1])	16.1 (24.1)	12.4 (13.4)	-1.8 (18.9)	9.1 (20.7)
Median	10.0	9.0	-4.0	4.5
Maximum	77	47	81	81
Total [2]	35	34	33	102
PID_4 (mm)				
Minimum	-18	-5	-43	-43
Average (SD [1])	23.1 (25.3)	20.9 (16.7)	-1.8 (20.7)	14.3 (23.8)
Median	18.0	19.0	-2.0	12.0
Maximum	78	74	82	82
Total [2]	35	34	33	102
PID_5 (mm)				
Minimum	-11	-27	-34	-34
Average (SD [1])	28.2 (23.3)	30.6 (22.2)	-2.4 (20.7)	19.1 (26.5)
Median	30.0	32.5	-7.0	19.0
Maximum	73	87	81	87
Total [2]	35	34	33	102
PID_6 (mm)				
Minimum	-7	-14	-32	-32
Average (SD [1])	35.7 (25.0)	37.4 (21.3)	-1.9 (19.8)	24.1 (28.5)
Median	40.0	38.0	-5.0	26.5
Maximum	82	89	83	89
Total [2]	35	34	33	102

Analyzed population: ITT Dataset: VasPid Data image: 2018-12-05

Comments:

1) SD = Standard deviation.
2) Number of analyzed data (without missing values).

7.1.7.2 Final Efficacy Analyses of Both Pads (PID6)

Table 6 PID6 – Analysis of variance (ITT population)

	Degrees of freedom[1)]	Sum of squares	Mean square	*F*-value	*p*-value
Treatment group	2	28467	14233.4	36.871	<.0001
Pain intensity at Visit 1	1	10986	10986.1	28.459	<.0001
Residues	98	37831	386.0		

Analyzed population: ITT Dataset: VasPid Data image: 2018-12-05

Comments:

1) Type III based on SAS terminology

Table 7 PID6 – Estimates of differences between treatment groups (ITT population)

	Estimate [1)]	SE [2)]	95% RI [3)]
Dvectis Single – Dvectis Double	-1.4	4.7	-10.8 - 8.0
Dvectis Single-control	35.2	4.8	25.7 - 44.7
Dvectis Double-control	36.6	4.8	27.0 - 46.2

Analyzed population: ITT Dataset: VasPid Data image: 2018-12-05

Comments:

1) Point estimate of difference.
2) Standard error of the estimate difference.
3) Final 95% interval of reliability for the difference in treatments, without correction to the comparison-wise error rate. For defined ordering and with the condition of stopping after the first interim analysis for success/efficacy being met, this interval is identical to the usual 95% reliability interval not taking interim into account. Compliance with the condition of early stopping is presented in the interim analysis results.

Table 8 PID$_6$ – Estimates of differences between treatment groups (PPS population)

	Estimate [1]	SE [2]	95% RI [3]
Dvectis Single – Dvectis Double	-1.4	5.1	-11.6 - 8.8
Dvectis Single-control	34.7	4.8	25.1 - 44.3
Dvectis Double-control	36.1	5.0	26.1 - 46.2

Analyzed population: PPS Dataset: VasPid Data image: 2018-12-05

Comments:

1) Point estimate of difference.
2) Standard error of the estimate difference.
3) Final 95% interval of reliability for the difference in treatments, without correction to the comparison-wise error rate. For defined ordering and with the condition of stopping after the first interim analysis for success/efficacy being met, this interval is identical to the usual 95% reliability interval not taking interim into account. Compliance with the condition of early stopping is presented in the interim analysis results.

7.1.7.3 PID$_6$ Exploratory Analysis

Table 9 Correlation coefficients and *p*-values for correlation between *PID$_6$* and the average time of pad use

	Dvectis Single		Dvectis Double	
	coefficient	*p*-value	coefficient	*p*-value
Pearson correlation coefficient r	-0.0716	0.6875	-0.1954	0.2759
Kendall correlation coefficient τ	0.0339	0.7779	-0.0758	0.5350
Spearman coefficient ρ	0.0268	0.8785	-0.1569	0.3754

Analyzed population: ITT Dataset: Correlation Data image: 2018-12-05

Comments:

The presented *p*-values relate to the null hypothesis of zero correlation between the quantities.

7.2 Secondary Quantities

No secondary quantities have been defined in the trial plan for efficacy.

7.3 Other Quantities

7.3.1 Analysis

Other quantities were processed within the exploratory analysis. The clinical trial was not planned in order to reject the potential hypotheses tested for these quantities with determined power. *Test p*-values can serve as indicators of interesting results worth further investigation; however, they cannot be used for drawing confirmatory conclusions.

The number of days when any analgesic treatment was used at least once was summarized and presented for the respective treatment groups. The use of nimesulide was summarized and presented in the same way.

The data recorded in the Oswestry questionnaire were evaluated in accordance with the scoring manual [*]. Aggregate statistics for Visit 2 and the difference between Visit 2 and Visit 1 were calculated and presented.

Strengthening of lumbar spine stabilizing muscles was summarized. The results were presented based on the treatment groups. The correlation between muscle strengthening and recorded total cumulative time of sitting on the pad was investigated in the pad groups. Correlation coefficients were calculated and presented (Pearson, Spearman and Kendall's tau) together with the null hypothesis test values *correlation coefficient=0*.

[*] The total score was evaluated by the investigator and filled in the patient's record sheet. Therefore the scoring manual procedure following was not a direct part of the analysis.

7.3.2 Method of Result Presentation

Table 10 Summaries for painkiller use

	Dvectis Single N = 35	Dvectis Double N = 34	No treatment N = 33	Total N = 102
Number of days when analgesic treatment was used				
Minimum	0	0	0	0
Average (SD [1])	3.7 (6.1)	4.0 (8.9)	14.7 (13.6)	7.4 (11.1)
Median	0.0	0.0	14.0	2.0
Maximum	25	42	42	42
Total [2]	34	33	32	99
Number of days when nimesulide was used				
Minimum	0	0	0	0
Average (SD [1])	0.8 (2.6)	1.6 (7.3)	8.0 (11.5)	3.4 (8.5)
Median	0.0	0.0	0.0	0.0
Maximum	13	42	42	42
Total [2]	34	33	32	99

Analyzed population: ITT Dataset: Diary Data image: 2018-12-05

Comments:

1) SD = Standard deviation.
2) Number of analyzed data (without missing values).

Table 11 Summaries of sonographic examinations

	Dvectis Single N = 35	Dvectis Double N = 34	No treatment N = 33	Total N = 102
Bederní m. multifidus vlevo, návštěva 1 (cm)				
Minimum	1.06	0.94	0.90	0.90
Average (SD [1])	1.527 (0.322)	1.484 (0.334)	1.678 (0.510)	1.561 (0.401)
Median	1.500	1.390	1.480	1.440
Maximum	2.50	2.50	2.78	2.78
Total [2]	35	34	33	102
Lumbar m. multifidus left, Visit 2 (cm)				
Minimum	1.10	1.03	0.92	0.92
Average (SD [1])	1.551 (0.323)	1.499 (0.310)	1.629 (0.476)	1.560 (0.376)
Median	1.500	1.415	1.485	1.460
Maximum	2.50	2.52	2.76	2.76
Total [2]	34	32	32	98
Lumbar m. multifidus left, change since Visit 1 at Visit 2 (cm)				
Minimum	-0.06	-0.03	-0.31	-0.31
Average (SD [1])	0.029 (0.031)	0.032 (0.025)	-0.033 (0.107)	0.010 (0.071)
Median	0.030	0.035	-0.005	0.020
Maximum	0.09	0.09	0.25	0.25
Total [2]	34	32	32	98
Lumbar m. multifidus right, Visit 1 (cm)				
Minimum	1.01	0.95	0.97	0.95
Average (SD [1])	1.526 (0.332)	1.483 (0.353)	1.684 (0.486)	1.563 (0.400)
Median	1.560	1.435	1.560	1.480
Maximum	2.30	2.60	2.89	2.89
Total [2]	35	34	33	102
Lumbar m. multifidus right, Visit 2 (cm)				
Minimum	1.06	1.02	0.94	0.94
Average (SD [1])	1.554 (0.320)	1.513 (0.347)	1.620 (0.462)	1.562 (0.379)
Median	1.525	1.450	1.465	1.460
Maximum	2.30	2.64	2.80	2.80
Total [2]	34	32	32	98
Lumbar m. multifidus right, change since Visit 1 at Visit 2 (cm)				
Minimum	-0.10	0.00	-0.34	-0.34
Average (SD [1])	0.034 (0.034)	0.038 (0.026)	-0.047 (0.084)	0.009 (0.066)
Median	0.040	0.040	-0.020	0.020
Maximum	0.09	0.09	0.06	0.09
Total [2]	34	32	32	98

Analyzed population: ITT Dataset: Sono Data image: 2018-12-05

Comments:

1) SD = Standard deviation.
2) Number of analyzed data (without missing values).

Table 12 Correlation coefficients and *p*-values for correlation between the changes in sonography results and average time of pad use

	Dvectis Single		**Dvectis Double**	
	coefficient	*p*-value	coefficient	*p*-value
Sono left (change)				
Pearson correlation coefficient r	**-0.1563**	**0.3773**	**0.0968**	**0.5980**
Kendall correlation coefficient τ	**-0.1034**	**0.3863**	**0.0827**	**0.5014**
Spearman coefficient ρ	**-0.1657**	**0.3415**	**0.0873**	**0.6276**
Sono right (change)				
Pearson correlation coefficient r	**-0.0674**	**0.7047**	**0.2108**	**0.2468**
Kendall correlation coefficient τ	**-0.1070**	**0.3697**	**0.1190**	**0.3343**
Spearman coefficient ρ	**-0.1645**	**0.3452**	**0.1620**	**0.3675**

Analyzed population: ITT Dataset: Correlation Data image: 2018-12-05

Comments:

The presented *p*-values relate to the null hypothesis of zero correlation between the quantities.

Table 13 Oswestry

	Dvectis Single N = 35	Dvectis Double N = 34	No treatment N = 33	Total N = 102
Oswestry score after Visit 1				
Minimum	6	4	0	0
Average (SD [1])	26.5 (13.4)	21.2 (11.6)	26.9 (13.6)	24.9 (13.1)
Median	22.0	19.0	26.0	22.0
Maximum	64	50	56	64
Total [2]	35	34	33	102
Oswestry score at Visit 2				
Minimum	4	2	2	2
Average (SD [1])	17.5 (10.4)	11.7 (9.0)	28.6 (14.1)	19.2 (13.2)
Median	16.0	9.0	28.0	18.0
Maximum	46	36	68	68
Total [2]	34	32	32	98
Oswestry score change at Visit 2 since Visit 1				
Minimum	-34	-44	-24	-44
Average (SD [1])	-8.0 (9.2)	-8.9 (8.9)	1.7 (13.0)	-5.1 (11.5)
Median	-8.0	-8.0	0.0	-5.0
Maximum	17	4	38	38
Total [2]	34	32	32	98

Analyzed population: ITT Dataset: Oswestry Data image: 2018-12-05

Comments:

1) SD = Standard deviation.
2) Number of analyzed data (without missing values).

8 Safety Analysis

8.1 Adverse Events and Effects

8.1.1 Analysis

Summaries of adverse event and effect occurrences in the predefined categories stated in CRF were calculated. The summaries included numbers and percentages of subjects with the occurrence of a given event/effect in the respective categories and in total individually for the treatment groups and in total.

8.1.2 Method of Result Presentation

Table 14 Safety summary

	Dvectis Single N = 35	Dvectis Double N = 33	No treatment N = 32	Total N = 100
Headache	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Up. r. inflammation[1)]	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Dizziness	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Hypertension	0 (0.0%)	0 (0.0%)	1 (3.1%)	1 (1.0%)
Dyspnea	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Diarrhoea	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Constipation	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Nausea	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Vomiting	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Flatulence	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Rash	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Other	3 (8.6%)	3 (9.1%)	1 (3.1%)	7 (7.0%)
Any	3 (8.6%)	3 (9.1%)	2 (6.3%)	8 (8.0%)

Analyzed population: SAF Dataset: AeList Data image: 2018-12-05

Comments:

The table includes numbers of subjects with the incidence of the above adverse events and effects, not the numbers of adverse events or effects

1) Up. r. inflammation = Upper respiratory inflammation

9 Interpretation and Summary

Demographic and other quantities were summarized before the start of the treatment for reasons of study population description (Table 2). Testing of differences between the groups was not the focus; there are no significant differences. Treatment regimen compliance was summarized (Table 3). The average time of sitting on the Dvectis Single pad is about 3,5 hours and is approximately 40 minutes longer than for Dvectis Double. The difference is probably not statistically significant; however, a formal test was not planned or performed.

The study was stopped early after the 1st sequential interim analysis based on compliance with the criteria for early study stopping for efficacy using Pocock's method in accordance with the study protocol plan. The sequential interim analysis that led to it is presented in a separate report.

Pain intensities and their differences from the initial value were summarized (Table 4 and Table 5). The treatment factor and pain value at Visit 1 have a statistically significant impact on the PID6 primary quantity values (Table 6).

95% reliability intervals were calculated for the differences between the treatment groups for the primary quantity (Table 7). For the difference between Dvectis Single - control, this interval does not include zero, therefore the result is consistent with the pad efficacy statement within evaluation of the early stopping criteria, i.e. it is statistically significant. The interval for the Dvectis Double - control difference does not include zero either. On the other hand, the interval for difference between the pads does include zero, and therefore is not significant.

The statistical significance stated in the previous paragraph relates to the significance level $\alpha = 0.05$ (and $\alpha = 0.025$ unilaterally, respectively). Its maintenance is guaranteed:

- Correct sequential interim analysis performance;
- Compliance with the early study stopping criterion after the first sequential interim analysis using Pocock's method due to the Dvectis Single pad efficacy;
- Consistent result for Dvectis Single during the final analysis;
- The result for Dvectis Double was performed after the compliance with the condition of null hypothesis rejection for Dvectis Single.

With regard to the use of the *stage-wise ordering (analysis time ordering / Fairbanks and Madsen ordering*) [1] and to the fact that there was early stopping after the first analysis, the presented reliability intervals are identical to the naive bilateral 95% reliability intervals for fixed selection size.

Robustness of the results was successfully confirmed by an alternative population analysis (Table 8). Therefore both pads can be claimed efficacious.

The exploratory analyses of correlations between the effect and time of pad use do not show statistically significant values (Table 9).

Other exploratory analyses are presented as summaries. They imply lower use of analgesics when the pads are used in comparison to control (Table 10) in particular, and possibly some other changes, probably without statistical significance.

Safety summary in terms of a comparison of the treatment groups is not conclusive due to low frequencies.

10 References

[1] A. A. Tsiatis, G. L. Rosner, and C. R. Mehta, “Exact confidence intervals following a group sequential test,” *Biometrics*, vol. 40, no. 3, pp. 797–803, Sep. 1984.

[2] C. Jennison and B. W. Turnbull, *Group sequential methods with applications to clinical trials*. Boca Raton: Chapman & Hall/CRC, 2000.

[3] S. S. Emerson and T. R. Fleming, “Parameter Estimation Following Group Sequential Hypothesis Testing,” *Biometrika*, vol. 77, no. 4, p. 875, Dec. 1990.

[4] S. S. Emerson and J. M. Kittelson, “A Computationally Simpler Algorithm for the UMVUE of a Normal Mean Following a Group Sequential Trial,” *Biometrics*, vol. 53, no. 1, p. 365, Mar. 1997.

[5] S. S. Emerson, “S+SeqTrial: Technical Overview.” University Washington, 2007.

[6] R. Marcus, E. Peritz, and K. R. Gabriel, “On Closed Testing Procedures with Special Reference to Ordered Analysis of Variance,” *Biometrika*, vol. 63, no. 3, p. 655, Dec. 1976.

[7] H. M. J. Hung, S.-J. Wang, and R. O’Neill, “Statistical Considerations for Testing Multiple Endpoints in Group Sequential or Adaptive Clinical Trials,” *J. Biopharm. Stat.*, vol. 17, no. 6, pp. 1201–1210, Nov. 2007.

[8] A. C. Tamhane, C. R. Mehta, and L. Liu, “Testing a Primary and a Secondary Endpoint in a Group Sequential Design,” *Biometrics*, vol. 66, no. 4, pp. 1174–1184, Dec. 2010.