

---

## 統計報告書

---

慢性腰痛患者における**Dvectis Single**および**Dvectis Double**パッドの有効性・安全性を「パッド不使用」との比較で評価するための臨床試験

プロトコルコード: **DVE-17**

試験対象器具	Dvectis Single Dvectis Double
適応症	慢性腰痛
設計	ランダム化比較臨床試験
研究開始	2018年4月9日
研究終了	2018年8月20日
研究責任者	MUDr. Martin Holinka 整形外科 Karvinská hornická nemocnice a.s. Zakladatelská 975/22 735 06 Karviná - Nové Město
契約者	DD - Hippero s.r.o. třída Kpt. Jaroše 1922/3 602 00 Brno
文書作成者	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.



.....

Signature

7. 12. 2018

.....

Date

## 目次

1 略号一覧	6
2 はじめに	7
2.1 本文書の趣旨と目的	7
2.2 試験、目的、仮説の概要	7
2.2.1 試験の概要	7
2.2.2 主目的、初期量、一次仮説	7
2.3 研究プロトコルの統計解析計画との比較で検知される変更	7
2.4 統計計画との比較で検知される変更	8
3 解析対象集団と被験者数	9
3.1 ARS集団	9
3.2 ITT集団	9
3.3 PPS集団	9
3.4 安全性解析対象集団	10
3.5 最終解析の解析対象集団定義方法	10
3.6 解析対象集団と患者病歴の概要	11
3.7 選択サイズ計算	11
4 欠測データ	13
5 逐次および最終解析	14
6 研究被験者の特徴、治療計画の遵守	15
6.1 治療開始前に測定した人口動態量、その他の数量	15
6.1.1 解析	15
6.1.2 結果の提示	15
6.2 治療計画の遵守	17
7 有効性	18
7.1 初期量	18
7.1.1 基本情報	18
7.1.2 順序付け	18
7.1.3 一次仮説に関係する最終解析 (Dvectis Singleパッド)	19
7.1.4 Dvectis Doubleパッドの有効性に関係する最終解析	19
7.1.5 初期量予備解析	20

7.1.6	その他の臨床試験目的の最終解析	20
7.1.7	結果の提示	20
7.1.7.1	初期量と関連導出量の総統計	20
7.1.7.2	両パッドの最終有効性解析 (PID6)	24
7.1.7.3	PID6予備解析	25
7.2	二次量	25
7.3	その他の数量	26
7.3.1	解析	26
7.3.2	結果の提示方法	27
8	安全性解析	31
8.1	有害事象と副作用	31
8.1.1	解析	31
8.1.2	結果の提示方法	31
9	解釈とまとめ	32
10	参考資料	34

## 1 略号一覧

ANOVA	Variation analysis 差異解析
ARS	All Randomised Subjects 全ランダム化被験者 ランダム化研究被験者全員で構成される解析対象集団
CRF	Case Report Form 症例報告書
H0	Null hypothesis 帰無仮説
HA	Alternative hypothesis 対立仮説
RI	Reliability interval 信頼区間
ITT	Intent-To-Treat 治療企図 治療企図原則に基づき最大観察数について定義された解析対象集団（すなわち、研究プロトコルから著しく逸脱する被験者も含まれる）
LOCF	Last Observation Carried Forward 直近の先行観測値による補填（データ補完法）
NOCB	Next Observation Carried Backward 欠測時点以降最初の観測値による補填（データ補完法）
MCAR	Missing Completely At Random 完全にランダムに欠測（欠測データ分布の種類）
PI	Pain Intensity 疼痛強度 過去48時間に感じた最も強い痛みで、VASに基づき記録されたもの
PID	Pain Intensity Difference 疼痛強度変化量（1週目と比較したPI差）
PPS	Per Protocol Set プロトコルに適合した被験者集団 研究プロトコルから著しく逸脱していない研究被験者で構成される解析対象集団
SAF	Safety Analysis Set 安全性解析対象集団 安全性評価のための解析対象集団
SAS	SASインスティテュートのソフトウェア製品
SD	Standard deviation 標準偏差
UMVUE	Uniform Minimum Variance Unbiased Estimator 一様最小分散不偏推定量
VAS	Visual analogue scale ビジュアル・アナログ・スケール
WCS	Worst Case Scenario 最悪シナリオ（データ補完法）
Z	Z-statistics Z統計量

## 2 はじめに

### 2.1 本文書の趣旨と目的

臨床試験データ統計解析結果の提示

### 2.2 試験、目的、仮説の概要

#### 2.2.1 試験の概要

試験は、単施設、オープン、ランダム化、3群間臨床試験であって、Karviná Miners病院 (Karvinská hornická nemocnice a.s.) 整形外科で実施された。慢性腰痛患者最大198人の参加を計画し、これらの患者を3つの治療群にランダムかつ均等に割り付けた。

1. Dvectis Single padパッド使用
2. Dvectis Double padパッド使用
3. パッド不使用

1回の逐次中間解析が臨床試験の中間に計画された。この中間解析で有効性を証明できたので、試験は早期終了した。

各患者の臨床試験参加期間は6週間 ( $\pm 5$ 日) と予定した。個々の調査来院と手順の概要、研究対象集団、その他の詳細情報については、研究プロトコルを参照。

計画に従い、評価被験者には来院1および来院2で過去48時間を感じた最も強い痛み (PI) を質問した。痛みは、研究者の監督のもとでCRFのビジュアル・アナログ・スケール (VAS) に被験者が記録した。評価被験者はさらに、痛みの強さを患者日誌に個人で毎日記録した。

#### 2.2.2 主目的、初期量、一次仮説

臨床試験の主たる目的は、慢性腰痛患者についてDvectis Singleパッド使用の有効性を「パッド不使用」と比較して評価することだった。

PIを初期量とした。

第2週と第6週 (PID6) のPI変化量 (PID) で評価し、検定に基づきDvectis Singleが優位となることを一次仮説とした。

### 2.3 研究プロトコルの統計解析計画との比較で検知される変更

研究計画とは異なり、臨床試験は単一の研究センターで実施された。研究プロトコルに含まれる計画とは異なり、研究センターのファクターは解析ではいかなるかたちでも説明されておらず、治療と研究センターとの相互作用の存在は解析されていない。同じ理由で、可能な場合小規模なセンターを組み合わせるといった計画も採用していない。

これら以外に1回目の逐次中間解析計画で検知される変更はない。

## 2.4 統計計画との比較で検知される変更

統計計画は統計解析の実施に先立ち策定され、正式な承認を受けた。計画には、統計解析実施と結果提示方法の技術的情報が含まれていた。

統計計画から逸脱していたのは、治療計画遵守評価の要約統計量に関するものだけである（表3「治療計画の遵守」を参照）。当初の計画は累積着座時間の評価をベースにしていたが、実際には平均時間を解析した。累積（または個人の）着座時間がデータシートに含まれておらず、したがって解析データに含まれていなかったためである。この点では統計解析計画のミスではあるが、実質的な影響はない。



### 3 解析対象集団と被験者数

統計解析に先立ち、研究プロトコルに沿って4つの集団を定義した。

1. 全ランダム化被験者からなる集団 (ARS)
2. 治療企図のある全患者の集団 (ITT)
3. 臨床試験計画から著しく逸脱しない患者の集団 (プロトコルに適合した被験者集団 – PPS)
4. 安全性解析集団

各解析対象集団は、そのメンバーのリストに基づいている。治療群はそれぞれの集団に割り付けられた患者個人で違っていてもおかしくなかったが（研究プロトコルと統計解析計画を参照）、そのようなケースにはならなかった（ランダム化の結果、事実上患者全員が治療群に入った）。したがって、実際の治療群は患者全員のランダム化治療群と同じである。これについては本文書ではこれ以上議論しない。

#### 3.1 ARS集団

ARS集団は、臨床試験に参加しランダム化された患者全員からなる。研究集団を説明するにあたっては、ARS集団を用い、来院1で確定した人口動態量やその他の数量をとりまとめた。

#### 3.2 ITT集団

このケースではITT集団はARS集団と同じである。計画では、PIデータが得られないARS患者、本研究に繰り返し参加している患者はITTには含まれないことになっていたが、このような状況は確認されなかった。ITT集団は、有効性評価のための一次解析対象集団である。

#### 3.3 PPS集団

統計計画には、臨床試験計画から著しく逸脱する患者を除き、PPSには全ITT患者を含めることになっていた。著しい逸脱には以下が該当する。

1. 1つ、または複数の参加基準を満たさない場合
2. 治療割付けにミスがあった場合
3. 無許可の併用薬または併用療法の使用が記録で確認された場合
4. PI0値が欠測している場合
5. PI値欠測が25%超の場合

6. 治療計画への著しい違反や臨床試験計画からの別の著しい逸脱がある場合  
観察報告書と不具合報告書からは以下のことが読み取れる。

1. 参加基準を満たさない研究被験者は登録されていない。
2. 不適切に治療割付けされた研究被験者は特定されていない。
3. 無許可の併用療法の使用の記録は確認されていない。
4. *PI0*値欠測はない。
5. 患者45、53および94は、*PI*値欠測数が25%超のため（83%）、研究を早期に終了した。
6. 以下の患者については最低限の패드着座時間が観察できず、治療計画に反していた（21、23、35、42、55、57、60、72、75、76、84、86、88）。

したがって、プロトコルから著しく逸脱した研究被験者は16人、全体の15.7%だった。

PPS集団は、有効性評価のための二次解析対象集団である。この評価のタスクは、臨床試験の主要な結果の堅牢性を実証することだった。そのため、*PI*での最終解析では検定群の有効性を対照群と比較し（ITTに加えて）、臨床試験の主要な結論を確認した。

### 3.4 安全性解析対象集団

計画では、来院1直後に連絡がとれなくなった被験者、および検定群のいずれかに（実際に）割付けられたもののパッドが支給されなかった被験者を除き、安全性解析対象集団はARSの被験者で構成されることになっていた。

パッドは検定群の患者全員に支給された。53と94の患者は連絡がとれなくなったため、安全性解析集団から除外された。

安全性解析対象集団の患者は、ランダムに割付けられていた群の如何にかかわらず、実際に割付けられた治療群に基づき解析された。

### 3.5 最終解析の解析対象集団定義方法

計画に従い解析を実施するにあたり、必要とされる解析対象集団を定義した。

### 3.6 解析対象集団と患者病歴の概要

表1 治療群と解析集団の被験者数

	Dvectis Single	Dvectis Double	治療なし	合計
ARS集団	35	34	33	102
ITT集団	35	34	33	102
PPS集団	29	25	32	86
安全性解析集団	35	33	32	100

### 3.7 選択サイズ計算

選択サイズの計算は、Dvectis Single群と対照群の初期量に示される有効性推定差に基づいており、これは15mmであった。期待値の不確実度をふまえ、また約10mmの差異でも現実的であるのは明らかで、臨床的にも重要であることを考慮し、当初期待値が過度に楽観的だとしても、この微小な差異を十分に検出できるように中間解析は計画された。

患者数の確定には以下の前提を用いた。

1. 一次仮説は、群の比較、すなわちDvectis Singleパッドと対照群の比較に基づく。そのため、Dvectis Double群は選択サイズでは直接的な役割は果たしていない。
2. 帰無仮説は、比較群間の差異ゼロに基づく。
3. 対立仮説は、Dvectis Singleパッドに有利な差異10mmと一致する。
4. 結果の分布は両比較群で同じであって有効性依存ではなく、標準偏差15mmと一致する。この値は過去の研究に基づき概算した。
5. 帰無仮説の検定は片側ではなく、有意水準は $\alpha = 0.025$ である。
6. 対立仮説に対応する差異を検出するには検定強度90%が必要である（これは臨床試験完了時のステージに関係しない）。
7. 多くても2回の逐次解析を等間隔で行う。1回目は研究被験者の予測最大数の中間で行う。
8. 1回目の逐次解析（中間解析）の後、上記のとおり、明確に定義された完了基準に基づき有効性（成功）が確認できれば、研究を終了することができる。
9. 検定統計の極限はポコック（Pocock）法に準拠し確定されている（いずれの逐次解析についても同じであり、（5）に示される既定の有意水準での第一種過誤発生の累積確率の維持を保証する）。

上記前提を導入した結果、各群に最大**52.02**人の評価可能被験者が必要となった。臨床試験のバランスの要件を満たすため（**Dvectis Double**パッド群も同じ被験者数である必要がある）、合計で **$3 \times 52.02 = 156.06$** 人の評価可能被験者が必要になった。

一次解析は解析対象ITT群で実施されたが、解析対象集団選択と臨床試験計画からの逸脱に対する結論の無感度の観点から、臨床試験の確実性を実証する必要があるため、評価可能要件は、重大な逸脱なしに分析対象の母集団に関連付ける必要があります。

臨床試験計画から著しく逸脱する被験者は**20%**未満であることを前提とすると、合計で約 **$156.06 / 0.8 = 195.08$** 人の被験者を臨床試験に参加させなければならなかった。**6**（**3**群×**2**）で割れる整数で**195.08**の次の整数は**198**である。したがって、各治療群で最大**66**人の被験者を臨床試験に含めなければならず、また中間解析は、**2**つの群（**Dvectis Single**群と対照群）の少ないほうの群で**33**人の被験者が完了した時点で実施しなければならなかった。

## 4 欠測データ

計画では、ARS集団解析の欠測データは欠測として処理されることとなっていた。すなわち、欠測データの合計数として要約統計量に記載されることになっていたが、欠測データは他の統計の計算から除外された。実際に有意であるのはパッド着座時間だけとなった。

PI量の全解析では、計画に従い、直近の先行観測値による補填 (LOCF) のデータ補完法を用いた。PI1が欠測する場合の計画は、そのような状況は生じなかったので使用しなかった。

LOCFを用いると一般的にはバイアスが生じるものだが、検定対象群の研究に患者が参加している期間、対照群と比較して観測されるPI値の減少に関しては、このようなアプローチは明らかに穏当である。

オスウェストリー (Oswestry) 質問票の評価については、主として個々の質問のレベルでLOCFの使用を計画し、また欠測値があれば、それを説明するためのスコアリングマニュアル手順も計画した。質問票は研究者が評価し、患者記録シートに合計スコアだけを記入しているので、この手順はデータ管理および統計解析活動の範囲外だった。

合計スコアが欠測している場合、ARS集団解析での欠測値処理と同様に処理した (すなわち解析からの排除)。

## 5 逐次および最終解析

1回目の逐次解析は中間解析計画に従って実施され、有効性が認められたので研究は早期終了とした。1回目の逐次解析の結果報告書は、別文書に含まれている。次に最終解析を実施し、これは本報告書に提示されている。

## 6 研究被験者の特徴、治療計画の遵守

逐次中間解析における研究集団の概要については、統計解析計画で確定した範囲で治療を開始する前に定められているとおり、人口動態量の要約統計量を計算し、提示した。要約統計量は各治療群（全3つの治療群）について提示してある。これらの特徴の差異の有意性は検定していない。

治療計画を遵守しているかは、研究に参加した期間における器具使用平均時間の総統計量を計算し評価した。

### 6.1 治療開始前に測定した人口動態量、その他の数量

#### 6.1.1 解析

解析した数量については、各治療群について合計として要約統計量を計算した。解析対象ARS集団を使用した。

#### 6.1.2 結果の提示

表2 治療開始前の人口動態その他の数量

	Dvectis Single N = 35	Dvectis Double N = 34	治療なし N = 33	合計 N = 102
<b>年齢 (年数)</b>				
最小値	29	26	37	26
平均 (SD <sup>1)</sup> )	50.1 (9.1)	45.9 (9.9)	51.1 (7.5)	49.0 (9.1)
中央値	52.0	47.0	52.0	50.0
最大値	65	60	63	65
合計 <sup>2)</sup>	35	34	33	102
<b>性別 N (%)</b>				
男性	3 (8.6%)	6 (17.6%)	7 (21.2%)	16 (15.7%)
女性	32 (91.4%)	28 (82.4%)	26 (78.8%)	86 (84.3%)
合計 <sup>2)</sup>	35	34	33	102
<b>体重 (kg)</b>				
最小値	43	52	53	43
平均 (SD <sup>1)</sup> )	78.7 (18.3)	78.4 (13.7)	78.3 (16.8)	78.5 (16.3)
中央値	75.0	78.5	78.0	77.5
最大値	120	103	123	123
合計 <sup>2)</sup>	35	34	33	102
<b>身長 (cm)</b>				
最小値	153	148	158	148
平均 (SD <sup>1)</sup> )	168.1 (8.3)	168.4 (8.0)	168.7 (7.3)	168.4 (7.8)
中央値	168.0	168.0	168.0	168.0
最大値	189	185	187	189
合計 <sup>2)</sup>	35	34	33	102
<b>治療開始前の疼痛強度 (mm VAS)</b>				
最小値	45	45	43	43
平均 (SD <sup>1)</sup> )	62.6 (14.2)	63.0 (13.7)	59.5 (12.8)	61.7 (13.5)
中央値	60.0	60.0	59.0	60.0
最大値	98	90	96	98
合計 <sup>2)</sup>	35	34	33	102

解析対象集団: ARS

データセット: ベースライ  
ン

データ画像: 2018年12月5日

**備考:**

1) SD = 標準偏差

2) 解析データ数 (欠測値なし)



## 6.2 治療計画の遵守

表3 治療計画の遵守

Dvectis Single	Dvectis Double	治療なし	合計	
N = 35	N = 34	N = 33	N = 102	
1日の平均パッド着座時間				
最小値	0.32	0.42	Not applicable	0.32
平均 (SD <sup>1)</sup> )	3.496 (2.253)	2.864 (1.472)	Not applicable	3.185 (1.921)
中央値	2.806	2.542	Not applicable	2.611
最大値	8.69	6.28	Not applicable	8.69
合計 <sup>2)</sup>	34	33	0	67

解析対象集団: ARS

データセット: 日誌

データ画像: 2018年12月5日

### コメント:

1) SD = 標準偏差

2) 解析データ数 (欠測値なし)

## 7 有効性

### 7.1 初期量

#### 7.1.1 基本情報

初期量は、100mmビジュアル・アナログ・スケール (VAS) で測定した疼痛強度 ( $PI$ ) である。

測定は第1週から第6週まで週1回行った ( $PI_1, PI_2, \dots, PI_6$ )。

各治療群の基本の総統計を計算し、各週個別に提示してある。

疼痛強度変化量 ( $PID$ ) は、相関関係に基づき第2週～第6週で  $PI_i$  から計算した。

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

各治療群の基本の総統計を計算し、各  $PID_i$  について提示してある。

一次仮説の検証には  $PID_6$  を使用し、臨床試験 (プロトコル) では次のように統計的に定式化される:  $\mu_{DS}$  を Dvectis Single の平均値  $PID_6$  とし、 $\mu_{Ref}$  を対照 (レファレンス) 群の平均値  $PID_6$  とすると、初期帰無仮説  $H_0$  は

$$H_0: \quad \mu_{DS} - \mu_{Ref} \leq 0$$

対立仮説  $H_A$  に対して、

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

有意水準  $\alpha = 0.025$  で片側検定される。

この検定とそれに続く有効性の結論は逐次解析結果に基づいているので、検定そのものは本文書に提示した最終解析には含まれない。

$PID_6$  の解析は、ANOVAモデルで実施した。このモデルには治療ファクターと  $PI_1$  共変量が含まれている。結果は、タイプIII平方和 (SAS用語) を含め、標準的なANOVA表の形式で提示した。

#### 7.1.2 順序付け

最終解析の実施にあたって順序付けは重要である。仮説の検定、 $p$ -値の計算、区間の構成のベースとなっているのは、結果の統計量を順序付けでき、順序付け基準は結果の極端性である一方、帰無仮説は有効であるという考え方である。例:  $Z$ -統計量は極端になるほど、その (絶対) 値は高くなる。結果は1つの数字ではなく一対の数字 ( $M, Z$ ) であり、また実施された逐次解析と計算された  $Z$ -統計量の範囲内で臨床試験が終了する場合は中間解析 (ステージ) の数字で構成されるため、逐次解析の実施は複雑になる。したがって、極端性による結果の順序付けの方法を定義しなければならなかった。

順序付けを定義することで、任意の  $M1$ 、 $M2$ 、 $Z1$ 、 $Z2$  の不等の確率について判断を下せる。しかしながら逐次統計の密度は単調値関係を持たないので、最適な順序付けの方法は一般的には確定できない。したがって、最高検出力での検定に関しては選択基準を適用できない。一般的には直感的選択肢がいくつかある。

研究プロトコルの計画および統計分析計画では、一次仮説に關係する結果はステージ基準順序付け（解析時間順序／フェアバンクスとマドセンの順序付け）[1]に基づいて順序付けされている。とくに次の理由から、多くの著者がこの選択肢を直感的に有効とみなしている。

- 最終解析で得られる  $P$ -値は、帰無仮説が棄却される場合にかぎり、中間解析（第1および第2ステージ後）の逐次検定で用いられる有意水準  $\alpha$  を下回る[2]。この特徴は結果の内部整合性との関連で重要であり、一般的には他の順序付けタイプにはみられないものである。
- 臨床試験が早期終了となる場合、 $P$ -値は情報量をベースにしない。あるいは試験が早期終了しない場合であれば、含まれる患者数をベースにしない[2]。

### 7.1.3 一次仮説に關係する最終解析（Dvectis Singleパッド）

臨床試験が逐次解析時に観察された極端な結果に基づき終了する場合、逐次解析なしで患者数固定の設計に共通してなされる推定値を転用する。詳細な情報は、研究プロトコルと統計計画に示されている。

このケースでは、臨床試験は1回目の逐次中間解析後に早期終了した。この事実に加え、上記の順序付けを適用することで、固定の選択サイズにナイーブベイズ法で計算した95%信頼区間と同じ推定値になった。

### 7.1.4 Dvectis Doubleパッドの有効性に關係する最終解析

Dvectis Doubleパッドとプラシーボの差異の検定は、臨床試験の重要な副次的目的であり、有効性についての追加的な表明を実現する。主目的を示す帰無仮説が臨床試験のいずれかのステージで棄却された場合にかぎり、差異有意性の検定を実施できることになっていた。この条件は1回目の逐次解析後に満たされた。二次仮説検定は1回のみ実施した。

一次帰無仮説が棄却された場合にかぎりという検定実施条件は一種の階層検定であり、前述の水準  $\alpha$  での第一種過誤発生 of 累積確率合計（ファミリーワイズ・エラー率）を維持しやすい。これは単に、いずれの検定も個別的に（階層的に）同じ水準  $\alpha$  で実施されることによるもので、閉検定原理を用いて正式に証明できる[6]。しかしながら、この手順は逐次中間解析にそのまま適用できないため、この臨床試験では過誤増大確率  $\alpha$  を厳密に管理することはできない。ハング（Hung）[7]は有意水準  $\alpha$  での二次仮説検定に基づく直感的な戦略を提案しており、これは合計有意水準  $\alpha$  で一次帰無仮説が棄却された場合、標準的なガウス分布の  $(1-\alpha)$  パーセンタイルに対応する（すなわち逐次検定内での合計）。

この戦略が第一種過誤の発生確率増大を厳密にコントロールできるとはかぎらないことは明らかだが、このケースの場合、深刻な問題とはならない[8]。

上記に関して、Dvectis Doubleパッドとプラシーボの差異は、検定群と対照群の差異に標準の両側95%信頼区間の計算により、有意水準=0.025を用い、0に対する位置を決め、片側優位検定に基づき解析した。この検定は帰無仮説を棄却し（「結果」を参照）、また（Dvectis Singleの）検定実施前に一次帰無仮説が棄却されているので、これが最終的な結果となる。

### 7.1.5 初期量予備解析

予備解析では、PID6と記録されたパッド着座時間合計週平均の相関をパッド群で調査した。相関係数を計算し、帰無仮説検定値相関係数=0の $p$ -値と併せて提示した（ピアソン相関係数、スピアマン順位相関係数 $\rho$ 、ケンドールの $\tau$ ）。しかし、これらの $p$ -値は更なる研究に値する興味深い結果の指標にしかない。そもそも、これらは確認的な結論の導出を目的とはしていない。

### 7.1.6 他の臨床試験目的の最終解析

確認目的のために計画された他の解析はなく、これらは実施された逐次解析の如何にかかわらず処理されている。

### 7.1.7 結果の提示

#### 7.1.7.1 初期量と関連導出量の総統計

表4 PI要約統計量

	Dvectis Single N = 35	Dvectis Double N = 34	治療なし N = 33	合計 N = 102
<b>PI<sub>1</sub> (mm)</b>				
最小値	45	45	43	43
平均 (SD <sup>1)</sup> )	62.6 (14.2)	63.0 (13.7)	59.5 (12.8)	61.7 (13.5)
中央値	60.0	60.0	59.0	60.0
最大値	98	90	96	98
合計 <sup>2)</sup>	35	34	33	102
<b>PI<sub>2</sub> (mm)</b>				
最小値	6	17	1	1
平均 (SD <sup>1)</sup> )	53.1 (23.9)	57.9 (13.5)	60.2 (19.4)	57.0 (19.5)
中央値	60.0	59.0	59.0	59.0
最大値	98	91	94	98
合計 <sup>2)</sup>	35	34	33	102
<b>PI<sub>3</sub> (mm)</b>				
最小値	2	26	4	2
平均 (SD <sup>1)</sup> )	46.4 (23.5)	50.6 (12.8)	61.4 (17.7)	52.7 (19.4)
中央値	52.0	50.5	60.0	55.0
最大値	98	77	95	98
合計 <sup>2)</sup>	35	34	33	102
<b>PI<sub>4</sub> (mm)</b>				
最小値	0	14	3	0
平均 (SD <sup>1)</sup> )	39.5 (24.6)	42.1 (14.7)	61.3 (20.1)	47.4 (22.3)
中央値	45.0	42.0	61.0	49.5
最大値	98	73	98	98
合計 <sup>2)</sup>	35	34	33	102
<b>PI<sub>5</sub> (mm)</b>				
最小値	0	3	4	0
平均 (SD <sup>1)</sup> )	34.4 (22.3)	32.4 (17.3)	61.9 (20.6)	42.6 (24.1)
中央値	30.0	28.5	65.0	42.5
最大値	98	72	95	98
合計 <sup>2)</sup>	35	34	33	102
<b>PI<sub>6</sub> (mm)</b>				
最大値	2	1	2	1
平均 (SD <sup>1)</sup> )	26.9 (23.5)	25.6 (16.9)	61.4 (18.2)	37.6 (25.6)

中央値	20.0	19.5	62.0	39.0
最大値	98	64	92	98
合計 <sup>2)</sup>	35	34	33	102

---

解析対象集団: ITT

データセット: VasPid

データ画像: 2018年12月5日

**Comments:**

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

表5 PID要約統計量

	Dvectis Single N = 35	Dvectis Double N = 34	治療なし N = 33	合計 N = 102
<b>PID<sub>2</sub> (mm)</b>				
最大値	-29	-20	-32	-32
平均 (SD <sup>1)</sup> )	9.5 (24.1)	5.1 (13.7)	-0.6 (19.7)	4.8 (19.9)
中央値	2.0	0.0	-2.0	0.5
最大値	73	40	84	84
合計 <sup>2)</sup>	35	34	33	102
<b>PID<sub>3</sub> (mm)</b>				
最小値	-24	-10	-25	-25
平均 (SD <sup>1)</sup> )	16.1 (24.1)	12.4 (13.4)	-1.8 (18.9)	9.1 (20.7)
中央値	10.0	9.0	-4.0	4.5
最大値	77	47	81	81
合計 <sup>2)</sup>	35	34	33	102
<b>PID<sub>4</sub> (mm)</b>				
最小値	-18	-5	-43	-43
平均 (SD <sup>1)</sup> )	23.1 (25.3)	20.9 (16.7)	-1.8 (20.7)	14.3 (23.8)
中央値	18.0	19.0	-2.0	12.0
最大値	78	74	82	82
合計 <sup>2)</sup>	35	34	33	102
<b>PID<sub>5</sub> (mm)</b>				
最小値	-11	-27	-34	-34
平均 (SD <sup>1)</sup> )	28.2 (23.3)	30.6 (22.2)	-2.4 (20.7)	19.1 (26.5)
中央値	30.0	32.5	-7.0	19.0
最大値	73	87	81	87
合計 <sup>2)</sup>	35	34	33	102
<b>PID<sub>6</sub> (mm)</b>				
最小値	-7	-14	-32	-32
平均 (SD <sup>1)</sup> )	35.7 (25.0)	37.4 (21.3)	-1.9 (19.8)	24.1 (28.5)
中央値	40.0	38.0	-5.0	26.5
最大値	82	89	83	89
合計 <sup>2)</sup>	35	34	33	102

解析対象集団: ITT

データセット: VasPid

データ画像: 2018年12月5日

## 備考:

1) SD = 標準偏差

2) 解析データ数 (欠測値なし)



## 7.1.7.2 両パッドの最終有効性解析 (PID6)

表6 PID<sub>6</sub> – 分散解析 (ITT集団)

	自由度 <sup>1)</sup>	平方和	平均平方	F-値	p-値
治療群	2	28467	14233.4	36.871	<.0001
来院1の疼痛強度	1	10986	10986.1	28.459	<.0001
留数	98	37831	386.0		
解析対象集団: ITT	データセット: ト: VasPid		データ画像:	2018年12 月5日	

## 備考:

1) SAS用語に基づくタイプIII

表7 PID<sub>6</sub> – 治療群間の差異推定 (ITT集団)

	推定 <sup>1)</sup>	SE <sup>2)</sup>	95% RI <sup>3)</sup>
Dvectis Single - Dvectis Double	-1.4	4.7	-10.8 - 8.0
Dvectis Single-対照	35.2	4.8	25.7 - 44.7
Dvectis Double-対照	36.6	4.8	27.0 - 46.2
解析対象集団: ITT	データセット: VasPid		データ画像: 2018年12月5日

## 備考:

- 1) 差異推定値
- 2) 推定差異の標準誤差
- 3) 治療での差異の最終95%区間、比較基準過誤率との相関なし。所定の順序付けから、また成功/有効性の1回目中間解析後終了条件が満たされているので、この区間は中間を考慮しない通常の95%信頼区間と同一である。早期終了条件を満たしていることは中間解析の結果に提示されている。

表8 PID<sub>6</sub> – 治療群間差異推定 (PPS集団)

	推定 <sup>1)</sup>	SE <sup>2)</sup>	95% RI <sup>3)</sup>
Dvectis Single - Dvectis Double	-1.4	5.1	-11.6 - 8.8
Dvectis Single-対照	34.7	4.8	25.1 - 44.3
Dvectis Double-対照	36.1	5.0	26.1 - 46.2

解析対象集団: PPS      データセット: VasPid      データ画像: 2018年12月5日

## 備考:

- 1) 差異推定値
- 2) 推定差異の標準過誤
- 3) 治療での差異の最終95%区間、比較基準過誤率との相関なし。所定の順序付けから、また成功/有効性の1回目中間解析後終了条件が満たされているので、この区間は中間を考慮しない通常の95%信頼区間と同一である。早期終了条件を満たしていることは中間解析の結果に提示されている。

7.1.7.3 PID<sub>6</sub> 予備解析表9 PID<sub>6</sub>と平均パッド使用時間の相関についての相関係数とp値

	Dvectis Single		Dvectis Double	
	係数	p-値	係数	p-値
ピアソン相関係数 $r$	-0.0716	0.6875	-0.1954	0.2759
ケンドール相関係数 $\tau$	0.0339	0.7779	-0.0758	0.5350
スピアマン係数 $\rho$	0.0268	0.8785	-0.1569	0.3754

解析対象集団: ITT      データセット: 相関      データ画像: 2018年12月5日

## 備考:

提示されているp値は、数量間ゼロ相関の帰無仮説に関する。

## 7.2 二次量

有効性に関する試験計画では二次量は定義されていない。

## 7.3 その他の数量

### 7.3.1 解析

その他の数量は予備解析で処理されている。この臨床試験は、これらの数量について所定の検出力で検定された潜在的仮説を棄却するために計画されたものではない。検定 $p$ -値は更なる研究に値する興味深い結果の指標とはなるものの、これらは確認的結論の導出には使用できない。

各治療群について、鎮痛療法が最低でも1回は用いられた場合の日数の要約統計量を出し、提示してある。ニメスリドの使用も同様に要約統計量を出し、提示してある。

オスウェストリー質問票に記録されたデータは、スコアリングマニュアルに従って評価されている\*。来院2の総統計、来院2と来院1の変化量を計算し、提示した。

腰椎安定化筋強化の概要も示している。結果は治療群に基づき提示した。筋肉強化と記録されたパッド着座累積時間の相関は、パッド群で調査した。相関係数を計算し、帰無仮説検定値相関係数=0と併せて提示した（ピアソン、スピアマン、ケンドールの $T$ ）。

---

\* 合計スコアは研究者が評価し、患者の記録シートに記入した。したがってスコアリングマニュアル手順を踏むことは解析に直接的には含まれない。

### 7.3.2 結果の提示方法

表10 鎮痛薬使用の要約統計量

	Dvectis Single N = 35	Dvectis Double N = 34	治療なし N = 33	合計 N = 102
<b>鎮痛療法使用日数</b>				
最小値	0	0	0	0
平均 (SD <sup>1)</sup> )	3.7 (6.1)	4.0 (8.9)	14.7 (13.6)	7.4 (11.1)
中央値	0.0	0.0	14.0	2.0
最大値	25	42	42	42
合計 <sup>2)</sup>	34	33	32	99
<b>ニメスリド使用日数</b>				
最小値	0	0	0	0
平均 (SD <sup>1)</sup> )	0.8 (2.6)	1.6 (7.3)	8.0 (11.5)	3.4 (8.5)
中央値	0.0	0.0	0.0	0.0
最大値	13	42	42	42
合計 <sup>2)</sup>	34	33	32	99

解析対象集団: ITT

データセット: 日誌

データ画像: 2018年12月5日

**備考:**

- 1) SD = 標準偏差
- 2) 解析データ数 (欠測値なし)

表11 音波検査の要約統計量

	Dvectis Single N = 35	Dvectis Double N = 34	治療なし N = 33	合計 N = 102
<b>腰部筋肉 左多裂筋、来院1 (cm)</b>				
最小値	1.06	0.94	0.90	0.90
平均 (SD <sup>1)</sup> )	1.527 (0.322)	1.484 (0.334)	1.678 (0.510)	1.561 (0.401)
中央値	1.500	1.390	1.480	1.440
最大値	2.50	2.50	2.78	2.78
合計 <sup>2)</sup>	35	34	33	102
<b>腰部筋肉 左多裂筋、来院2 (cm)</b>				
最小値	1.10	1.03	0.92	0.92
平均 (SD <sup>1)</sup> )	1.551 (0.323)	1.499 (0.310)	1.629 (0.476)	1.560 (0.376)
中央値	1.500	1.415	1.485	1.460
最大値	2.50	2.52	2.76	2.76
合計 <sup>2)</sup>	34	32	32	98
<b>腰部筋肉 左多裂筋、来院2での来院1以降の変化 (cm)</b>				
最小値	-0.06	-0.03	-0.31	-0.31
平均 (SD <sup>1)</sup> )	0.029 (0.031)	0.032 (0.025)	-0.033 (0.107)	0.010 (0.071)
中央値	0.030	0.035	-0.005	0.020
最大値	0.09	0.09	0.25	0.25
合計 <sup>2)</sup>	34	32	32	98
<b>腰部筋肉 右多裂筋、来院1 (cm)</b>				
最小値	1.01	0.95	0.97	0.95
平均 (SD <sup>1)</sup> )	1.526 (0.332)	1.483 (0.353)	1.684 (0.486)	1.563 (0.400)
中央値	1.560	1.435	1.560	1.480
最大値	2.30	2.60	2.89	2.89
合計 <sup>2)</sup>	35	34	33	102
<b>腰部筋肉 右多裂筋、来院2 (cm)</b>				
最小値	1.06	1.02	0.94	0.94
平均 (SD <sup>1)</sup> )	1.554 (0.320)	1.513 (0.347)	1.620 (0.462)	1.562 (0.379)
中央値	1.525	1.450	1.465	1.460
最大値	2.30	2.64	2.80	2.80
合計 <sup>2)</sup>	34	32	32	98
<b>腰部筋肉 右多裂筋、来院2での来院1以降の変化 (cm)</b>				
最小値	-0.10	0.00	-0.34	-0.34
平均 (SD <sup>1)</sup> )	0.034 (0.034)	0.038 (0.026)	-0.047 (0.084)	0.009 (0.066)
中央値	0.040	0.040	-0.020	0.020
最大値	0.09	0.09	0.06	0.09
合計 <sup>2)</sup>	34	32	32	98

## 備考:

- 1) SD = 標準偏差
- 2) 解析データ数 (欠測値なし)

表12 音波検査結果の変化と平均パッド使用時間の相関についての相関係数と $p$ -値

	Dvectis Single		Dvectis Double	
	係数	$p$ -値	係数	$p$ -値
<b>音波検査 左 (変化)</b>				
ピアソン相関係数 $r$	-0.1563	0.3773	0.0968	0.5980
ケンドール相関係数 $\tau$	-0.1034	0.3863	0.0827	0.5014
スピアマン係数 $\rho$	-0.1657	0.3415	0.0873	0.6276
<b>音波検査 右 (変化)</b>				
ピアソン相関係数 $r$	-0.0674	0.7047	0.2108	0.2468
ケンドール相関係数 $\tau$	-0.1070	0.3697	0.1190	0.3343
スピアマン係数 $\rho$	-0.1645	0.3452	0.1620	0.3675

解析対象集団: ITT

データセット: 相関

データ画像: 2018年12月5日

## 備考:

提示されている $p$ -値は、数量間ゼロ相関の帰無仮説に関する。

表13 オスウェストリー

	Dvectis Single N = 35	Dvectis Double N = 34	治療なし N = 33	合計 N = 102
<b>来院1以降のオスウェストリースコア</b>				
最小値	6	4	0	0
平均 (SD <sup>1)</sup> )	26.5 (13.4)	21.2 (11.6)	26.9 (13.6)	24.9 (13.1)
中央値	22.0	19.0	26.0	22.0
最大値	64	50	56	64
合計 <sup>2)</sup>	35	34	33	102
<b>来院2時点のオスウェストリースコア</b>				
最小値	4	2	2	2
平均 (SD <sup>1)</sup> )	17.5 (10.4)	11.7 (9.0)	28.6 (14.1)	19.2 (13.2)
中央値	16.0	9.0	28.0	18.0
最大値	46	36	68	68
合計 <sup>2)</sup>	34	32	32	98
<b>来院2時点の来院1以降のオスウェストリースコアの変化</b>				
最小値	-34	-44	-24	-44
平均 (SD <sup>1)</sup> )	-8.0 (9.2)	-8.9 (8.9)	1.7 (13.0)	-5.1 (11.5)
中央値	-8.0	-8.0	0.0	-5.0
最大値	17	4	38	38
合計 <sup>2)</sup>	34	32	32	98

解析対象集団: ITT

データセット: オスウェストリー

データ画像: 2018年12月5日

**備考:**

- 1) SD = 標準偏差
- 2) 解析データ数 (欠測値なし)

## 8 安全性解析

### 8.1 有害事象と副作用

#### 8.1.1 解析

CRFに示される既定のカテゴリーにおける有害事象と副作用の発生の要約統計量を計算した。要約統計量には、各カテゴリーにおける所定の事象／作用が生じた被験者の人数とパーセンテージ、および各治療群の合計と全体の合計が含まれている。

#### 8.1.2 結果の提示方法

表14 安全性要約統計量

	Dvectis Single N = 35	Dvectis Double N = 33	治療なし N = 32	合計 N = 100
頭痛	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
上気道炎	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
めまい	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
高血圧	0 (0.0%)	0 (0.0%)	1 (3.1%)	1 (1.0%)
呼吸困難	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
下痢	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
便秘	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
吐き気	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
嘔吐	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
腹部の膨張	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
発疹	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
その他	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%)

解析対象集団: SAF      データセット: AeList      データ画像: 2018年12月5日

**備考:**

この表に記載される数字は、上記の有害事象や副作用を経験した被験者の人数であって、有害事象や副作用の件数ではない。



## 9 解釈とまとめ

人口動態その他の数量は、研究集団の概要を示す必要があったため、治療開始前にとりまとめた（表2）。各群間の差異を検定することは焦点ではなく、有意な差異はない。治療計画の遵守をとりまとめた（表3）。Dvectis Singleパッド着座平均時間は約3.5時間で、Dvectis Doubleよりも約40分長い。この差異は統計的に有意ではないと考えられるが、正式な検定は計画も実施もしていない。

研究プロトコル計画に従い、ポコックの方法に準拠し有効性が確認できた時点で研究を早期終了できるという基準を満たしていたので、研究は1回目の逐次中間解析後に早期終了した。この判断に至った逐次中間解析は別途の報告書に示されている。

疼痛強度と初期値からの変化量をとりまとめた（表4および表5）。来院1での治療ファクターと疼痛値は、PID6初期量の値に統計的に有意な影響を与える（表6）。

初期量について治療群間差異が計算され、95%信頼区間が算出された（表7）。Dvectis Singleと対照群の差異については、この区間にゼロは含まれないため、結果は早期終了基準の評価の範囲内でパッドの有効性表示と整合する（すなわち統計的に有意である）。Dvectis Doubleと対照群の差異に関する区間にもゼロは含まれない。一方、パッド間の差異についての区間にはゼロが含まれるので、有意ではない。

上記パラグラフにある統計的有意性は、有意水準 $\alpha=0.05$ （片側につき $\alpha=0.025$ ）を指す。この有意水準の維持は以下により保証される。

- 正しい逐次中間解析を実施している。
- ポコックの方法に準拠し、Dvectis Singleパッドの有効性の確認により、1回目の逐次中間解析後の研究早期終了基準を満たしている。
- 最終解析でDvectis Singleの結果が整合している。
- Dvectis Doubleの結果は、Dvectis Singleの帰無仮説棄却条件を満たした後で出されている。

ステージ基準の順序付け（解析時間順序／フェアバンクスおよびマドセンの順序付け）の使用 [1] および1回目の解析後に早期終了したことにに関して、提示されている信頼区間は、固定選択サイズの場合の両側95%信頼区間（ナイーブベイズ法）と同じである。

結果の堅牢性は、対立集団解析で完全に確認されている（表8）。したがって、どちらのパッドについても有効性を主張することができる。

パッド使用の効果と時間の相関に関する予備解析は統計的に有意な値は示していない（表9）。

その他の予備解析は、要約統計量として提示されている。これらは、とくに対照群との比較でパッド使用時に鎮痛薬の使用が減少することを示唆しており（表10）、おそらく他の変化もあるとは考えられるが、統計的有意性はないものと思われる。

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.