T2001-01 – Statistical Analysis Plan

An Observational Study to Collect Data Characterizing Analgesia in Patients Suffering from Bone Metastasis induced Pain.

DNAlytics contact

DNAlytics S.A. (referred to as "DNAlytics"), a Belgian-law company, with offices located Chemin du Cyclotron 6 in B-1348 Louvain-la-Neuve, Belgium, KBO/BCE 0844.973.235.

Dr Thibault Helleputte, CEO
Email: thibault.helleputte@dnalytics.com
Phone: +32 10 390096
Mobile: +32 498 642463

Ir Pierre Gramme, Principal Analyst
Email: pierre.gramme@dnalytics.com
Phone: +32 10 390096
Mobile: +32 498 616327

Dr Jérôme Paul, Analyst
Email: jerome.paul@dnalytics.com
Phone: +32 10 390096
Mobile: +32 473 459959

Document approval form

Document: T2001-01 – Statistical Analysis Plan

Study alias	T2001-01
Version of document submission	V1.3 - 2017-04-10

Document prepared by				
Company Name	**Name**	**Function**	**Signature**	**Date**
DNAlytics	Pierre Gramme	Principal Analyst		24/4/2017
DNAlytics	Thibault Helleputte	CEO		21/04/2017

Document approved by				
Company Name	**Name**	**Function**	**Signature**	**Date**
Tools4Patient	Alvaro Pereira	CSO		

Contents

Chemin du Cyclotron, 6 – B-1348 – Louvain-la-Neuve – Belgium
tva: be 0844.973.235 – iban: be29 5230 8050 8964 – bic: TRIOBEBB
DNAlytics SA is a Spin-Off Company of the Université catholique de Louvain, ICTEAM/INGI

1 Introduction

This statistical analysis plan covers the interim and final analyses of the Oncopain study T2001-01. As specified in the study protocol, the interim analysis will be executed as soon as 60 patients have completed the study.

The primary objective of the complete study consists in finding predictive factors of pain management and analgesic treatment efficacy in patients suffering from CIBP (cancer-induced bone pain). The analyses necessary to achieve this objective are very complex to conduct, due to the complexity of the data available. Secondary objectives describe intermediate milestones which will be reached during these analyses.

An interim analysis is also planned for the validation of sample size and necessary items among the personality tests.

All statistical analyses related to the study will be the responsibility of DNAlytics or designee. They will be conducted on the full analysis set. This set includes all data from patients who have been screened and who completed the study (no drop-outs). No safety analyses will be conducted during the interim study. The safety analyses on the complete database will be performed by the sponsor and reported in the Clinical Study Report. Additional exploratory analyses of the data will be conducted as deemed appropriate.

All hypothesis tests will be conducted at a two-sided alpha level of 0.05, unless otherwise stated.

1.1 Responsibilities

Name	Company	Role
Pierre Gramme	DNAlytics	Study design, analysis implementation, reporting
Jérôme Paul	DNAlytics	Study design, analysis review, reporting
Thibault Helleputte	DNAlytics	Study design, analysis and report review
Alvaro Pereira	Tools4Patient	Study design. As study sponsor representative, monitors data analysis progress.

1.2 Changes from Protocol

- A secondary endpoint was added for the analgesic treatment efficacy: Patient's daily APS and the linear regression coefficient of its daily evolution
- Secondary objectives of the final study were precised, while they were absent from the protocol

1.3 Amendment Rationale

The primary study objective is the prediction of individual analgesic treatment efficacy. All primary and secondary endpoints concerning treatment efficacy were based on differences between weekly average of pain measures at baseline and at the end of the follow-up period. Due to frequent changes in the analgesic therapy of cancer patients with bone metastases, it is essential to have a temporal granularity finer than the 4 weeks (i.e. follow-up duration). Therefore an additional endpoint was added for describing the daily values of APS and its linear trend throughout follow-up.

Chemin du Cyclotron, 6 – B-1348 – Louvain-la-Neuve – Belgium
tva: be 0844.973.235 – iban: be29 5230 8050 8964 – bic: TRIOBEBB
DNAlytics SA is a Spin-Off Company of the Université catholique de Louvain, ICTEAM/INGI

No secondary objective was described in the study protocol. For the sake of clarity, the current analysis plan describes 3 secondary objectives, all of which can be considered as intermediary milestones necessary in order to reach the (unmodified) primary objective of the study.

2 Study objectives

2.1 Primary Objective (for the whole study)

To define predictive factors of pain management/treatments efficacy in cancer patients with bone metastases suffering from pain (or cancer-induced bone pain (CIBP)), based on individual patient disease, therapy history, personality traits and expectation.

2.2 Secondary Objectives

1. To identify groups of patients showing similar profile at baseline, in terms of personality traits and disease history
2. To measure the concordance between questionnaire-based and clinical pain assessment
3. To identify typical evolution of pain level among cancer patients with bone metastases

2.3 Interim Objectives

1. To confirm the considered sample size of approximately 140 patients completing this study
2. To confirm questionnaires and historical data to be collected

2.4 Primary Endpoint

Individual analgesic treatment efficacy endpoint: patient's change from baseline of pain severity, as measured by the weekly means of the daily Average Pain Scores (APS).

2.5 Secondary Endpoints

- Individual analgesic treatment efficacy: Patient's daily APS and the linear regression coefficient of its daily evolution
- Individual analgesic treatment efficacy: Patient's change from baseline of
  - Pain severity, as measured by the weekly means of the daily Worst Pain Score (WPS),
  - Pain severity and interference as measured by the Brief Pain Inventory (BPI),
  - Investigator and Patient Global Assessment of Changes (IGAC and PGAC),
  - Pain intensity measured after Clinical Pain Assessments (CPAs) at Visit 2,
  - Quality-of-Life Questionnaire in 30 questions (QLQ-C30).
- Correlation between pain severity and clinical pain assessment as measured by:
  - Pain severity, as measured by the weekly Average Pain Score (APS) at Visit 1 and 2,
  - Pain intensity measured after Clinical Pain Assessments (CPAs) at Visit 1 and 2.

Variables used as predictive factors rather than as endpoints are descibed hereunder in Section 4.2.

Chemin du Cyclotron, 6 – B-1348 – Louvain-la-Neuve – Belgium
tva: be 0844.973.235 – iban: be29 5230 8050 8964 – bic: TRIOBEBB
DNAlytics SA is a Spin-Off Company of the Université catholique de Louvain, ICTEAM/INGI

3 Summary of Study Design

See the latest protocol version.

4 Populations for Analyses

Statistical analysis and data tabulation will be performed using the population of completers unless specified otherwise. This set includes all data from patients who have been screened, included and who completed the study (no drop-outs).

For the interim analysis, only patients who completed the study before the interim database transfer will be included.

4.1 Study Participant Disposition

All patients who discontinue from the study will be identified, and the extent of their participation in the study will be reported by the sponsor in the Clinical Study Report. If known, a reason for their discontinuation will be given.

4.2 Study Participant Characteristics

The following variables will be recorded and may be used as quantitative or classification variables:

- Patient characteristics, measured at V1:
  - Study center
  - Patient's sociodemographic characteristics: age, sex and other demographic characteristics
  - Recent analgesic treatment history
  - Personality questionnaires, at facet- and domain-level: Multidimensional Health Locus of Control Scale (MHLCS, with domains I, C and P), Hospital Anxiety and Depression Scale (HADS), Revised Life Orientation Test (LOT-R) and Expectation.
  - Pain evaluation questionnaires: Douleur Neuropathique en 4 questions (DN4), Neuropathic Pain Symptom Inventory (NPSI), and Pain Catastrophizing Scale (PCS)
- Pain and quality of life assessment, measured at V1 and V2:
  - Brief Pain Inventory (BPI)
  - Quality of Life questionnaire in 30 questions (QLQ-C30), IGAC, PGAC
  - Clinical Pain Assessment (CPA): Static Allodynia (SA), Dynamic Mechanical Allodynia (DMA), Mechanical Pain Threshold (MPT), Thermal Pain Sensitivity (TPS)
- Follow-up from the daily diary and intermediate visit:
  - Daily Average and Worst Pain Score (APS and WPS),
  - Daily analgesic therapy (including rescue medication): number of distinct drugs intaken, number of distinct drug categories intaken
  - Intake of rescue analgesic treatment if any,

Note that the description of analgesic therapy hereabove requires the grouping of drugs into a small number of classes which will be described in the Statistical Analysis Report.

The collection of the values of all these variables and variables derived from them for a single patient is referred to as *the patient-vector* below in the text.

Chemin du Cyclotron, 6 – B-1348 – Louvain-la-Neuve – Belgium
tva: be 0844.973.235 – iban: be29 5230 8050 8964 – bic: TRIOBEBB
DNAlytics SA is a Spin-Off Company of the Université catholique de Louvain, ICTEAM/INGI

5 Description of the Analyses

The final statistical analysis covers four aspects (Ax) detailed in the following sections:

- A1: Descriptive statistics at baseline, covering secondary objectives number 1 and 2.
- A2: Evolution of daily pain and treatments, covering secondary objective number 3.
- A3: Prediction of treatment efficacy, covering the primary objective

In addition, the interim statistical analysis will cover the following two aspects, corresponding to the interim objectives:

- I1: Sample size computation for the prediction of treatment efficacy.
- I2: Relation between psychological traits and pain evolution

5.1 A1: Descriptive statistics at baseline

This section provides general information useful before starting any data analysis. Only variables at baseline are analysed at this point, so that any decisions taken are not influenced by the observed response to treatment. The variables covered are: sociodemographic characteristics, recent analgesic treatment history, personality questionnaires, pain (including CPA) and quality of life at visit 1.

5.1.1 A1.1: Data preprocessing

Some variables may require preprocessing before being analysed. These preprocessing steps will be described here. It typically covers regrouping rare values of some categorial variables, applying logarithm transformation of some continuous variables, etc.

5.1.2 A1.2: Univariate statistics

Classical statistical information will be computed for each variable among the patient characteristics at V1. They will be listed and summarized in Table (tab-v1-summary). Descriptive statistics including subject count (N), mean, standard deviation (SD), median, first and third quartiles will be reported for continuous variables. Frequencies will be reported for categorical variables in Table (tab-v1-summary-mod). The number of missing values (NA) will also be reported for each variable. In subsequent analyses, value imputation might be performed if needed. Such cases will be documented.

5.1.3 A1.3: Outlier detection and clustering

Population outliers will be detected through a principal component analysis (PCA) visualization, reported in Figure (1). The percentage of variance captured will be documented. The patients will be colored by their study center in order to detect possible differences among the centers.

Outliers might be removed for at least parts of the further analyses. Such cases will be documented.

A clustering procedure, namely an ensemble k-Means, will be applied in order to identify subgroups of patients with intra-group similar variables at baseline, and inter-group different variables at baseline. Table (tab-v1-clustering-desc) will report per-cluster univariate statistics (median, first and third quartiles) for all variables used in the clustering procedure.

Chemin du Cyclotron, 6 – B-1348 – Louvain-la-Neuve – Belgium
tva: be 0844.973.235 – iban: be29 5230 8050 8964 – bic: TRIOBEBB
DNAlytics SA is a Spin-Off Company of the Université catholique de Louvain, ICTEAM/INGI

5.1.4 A1.4: Correlation analysis

The correlation between each variable (personality traits and sociodemographic variables) and the pain at baseline (APS, WPS, BPI severity, interference and QLQ-C30) will be assessed. Statistical tests for non-zero Spearman's correlation will be used. For categorial variables, exact Fisher tests will be used. Correction for multiple testing will be applied. Results will be reported in Table (tab-v1-correlations) and Figure (2). Scatterplots of the significant correlations will be produced. Figure (3) will detail the correlations within pain assessment variables and Figure (4) the correlations between personality traits and pain assessment.

5.2 A2: Evolution of daily pain and treatments

Pain evolution will first be analysed on its own (secondary objective 3, i.e. without linking it to the evolution of treatments). Then, the evolution of analgesic therapies are studied, as well as their correlation with pain evolution. Due to the complexity of treatments combinations for each patient and their frequent adaptation during the follow-up period, this analysis will be made on a simple summary of the therapy evolution. Finally, the interest of each variable characterizing the patient at baseline is evaluated for its relation with the pain evolution profile.

5.2.1 A2.1: Summary of individual pain evolution

Feature extraction will be performed in order to summarize each per-patient APS evolution curve into 3 meaningful variables:

- mean APS value
- linear trend of APS
- variability of APS around trend

The linear approximation of APS was chosen in order to describe the globally increasing or decreasing trend, while keeping easily interpretable results and avoiding overfitting. How to measure these values for one patient p:

- consider a model $APS_p(t) = a_{p0} + a_{p1}\ t$ where p is one patient and t is the time (in days since inclusion of patient p)
- Fit weights a_{pi} by linear regression on the patient's data (one dataset per patient, one record per timepoint)
- The variability around trend is measured by the per-patient RMSE of the linear model

Univariate statistics (mean, standard deviation (SD), median, first and third quartiles) on the distribution of these variables will be reported in Table (tab-pain-evol-distr). A one-sample Wilcoxon signed rank test will be performed in order to determine whether the linear trend in APS is globally different from zero. In addition, visualisation of the patients based on these 3 features will be performed in Figure (5).

5.2.2 A2.2: Summary of individual analgesic treatment evolution

Based on the description of analgesic therapy in the patient diary, the daily number of distinct drugs will be computed as well as the daily number of distinct drug categories. Treatment doses will be ignored for this analysis. The same feature extraction as in A2.1 will be performed on these daily numbers of drugs in order to obtain a per-patient average and linear trend of the number of drugs.

Univariate statistics (mean, standard deviation (SD), median, first and third quartiles) on the distribution of these variables will be reported in Table (tab-treat-evol-distr). A one-sample Wilcoxon signed rank test will be performed in order to determine whether the linear trend in number of analgesics is globally different from zero. In addition, visualisation of the patients based on these features will be performed in Figure (6).

Chemin du Cyclotron, 6 – B-1348 – Louvain-la-Neuve – Belgium
tva: be 0844.973.235 – iban: be29 5230 8050 8964 – bic: TRIOBEBB
DNAlytics SA is a Spin-Off Company of the Université catholique de Louvain, ICTEAM/INGI

5.2.3 A2.3: Relation between patient's characteristics and pain evolution

For each feature from the patient characteristics at V1, the correlation will be computed with the linear trend of pain evolution defined in A2.1. A p-value for non-zero correlation will be reported too, with BH correction for multiple testing. Results will be presented in Table (tab-cor-baseline-pain-evol).

5.3 A3: Prediction of treatment efficacy

A predictive model will be tuned based on the characteristics at V1 in order to predict the efficacy of each treatment on the patient. One model will be built for each treatment t, with the following design:

- Scope: patients having received treatment t on some but not all days
- Response (target): Per-patient treatment effect, as measured by the average difference of daily APS between the days with and without treatment t
- Predictors: patient characteristics at baseline

The performances of these pre-treatment models will be evaluated inside a resampling procedure. This procedure repeatedly samples 80% data to learn a predictive model whose predictive performances are assessed on the 20% remaining unseen data. In total, 200 train-test splits are performed for each treatment. The predictive model in use will be a ridge regression, as implemented by `glmnet::cv.glmnet` in R, with parameter `alpha=0`.

Table (tab-pred-perf) will report the average predictive performance on the test sets for each treatment, in terms of relevant statistics.

5.4 Interim analyses

The interim analyses will be performed on all patients having completed the study before the date of the interim database extract.

5.4.1 I1: Sample size computation for the prediction of treatment efficacy

For computing the necessary sample size, a *learning curve* will be evaluated for the analysis in A3. This analysis will first be run on all patients available in the interim database extract, and the average per-treatment performance will be computed. Then a random sample of 90% of the patients will be drawn and the same analysis will be repeated. And similarly with decreasing sample size.

This will lead to a "learning curve" showing the variation of predictive performance wrt to sample size and shown in Figure (7). An extrapolation of the curve will be performed in order to estimate the expected performance obtained in the final analysis using all 140 samples. A confirmation of infirmation of the number of patients to be included in the study will be made based on this estimation

5.4.2 I2: Relation between psychological traits and pain evolution

For each feature from the patient characteristics at V1, the correlation will be computed with the linear trend of pain evolution defined in A2.1. A p-value for non-zero correlation will be reported too, with BH correction for multiple testing. Results will be presented in Table (tab-cor-baseline-pain-evol).

Chemin du Cyclotron, 6 – B-1348 – Louvain-la-Neuve – Belgium
tva: be 0844.973.235 – iban: be29 5230 8050 8964 – bic: TRIOBEBB
DNAlytics SA is a Spin-Off Company of the Université catholique de Louvain, ICTEAM/INGI

6 Tables, Figures and Listings requirements

6.1 Tables

All tables resulting from the analyses will be grouped in an Excel file with one worksheet per table. The worksheet names correspond to the table name mentioned in the text above. Legends for all tables will be included on top of the corresponding worksheet. The table templates used for the final results will match attached file `T2001-01-SAP-table-templates-V1.xlsx`

6.2 Figures

Figure 1: Outlier detection by Principal Component Analysis based on data at baseline. The two first components are represented (Variance captured: %). Colors indicate the grouping of points into clusters.

Figure 2: Matrix of correlations between all pairs of variables at baseline (personality traits and sociodemographic variables)

Figure 3: Scatterplots of the significant correlations between the different pain measures at baseline

Figure 4: Scatterplots of the significant correlations between a personality trait and a pain measure at baseline

Figure 5: Summary visualization of the evolution of pain during follow-up period. Each point represents a patient, with average APS as abscissa and linear trend as ordinate

Figure 6: Summary visualization of the evolution of analgesic treatments during follow-up period. Each point represents a patient, with the number of different analgesic drugs (for the graph on the left) or the number of different analgesic drug classes (on the right) as abscissa, and their linear trend as ordinate

Figure 7: Learning curve for the prediction of analgesic treatment efficacy. The average performance evaluated by cross-validation on random subsets of the data is shown in ordinate, with the size of the training set as abscissa. An extrapolation curve (LOWESS) is also shown for sample sizes up to N=200.

7 Data Transfer and Unblinding Plan

The study data is collected and managed by Aepodia SA. Several information transfers will occur between Aepodia and DNAlytics, under the supervision of Tools4Patient:

Jun. 2015 **Database table structure transfer** – The objective of this transfer is to allow DNAlytics to prepare the analysis scripts and finalize the SAP. The data concerning a few included patients will be transfered in order to make the scripts creation possible. Those patients data are anonymized by Aepodia prior to the transfer to DNAlytics. Those patients will however be identified by their inclusion number.

Chemin du Cyclotron, 6 – B-1348 – Louvain-la-Neuve – Belgium
tva: be 0844.973.235 – iban: be29 5230 8050 8964 – bic: TRIOBEBB
DNAlytics SA is a Spin-Off Company of the Université catholique de Louvain, ICTEAM/INGI

May 2017 (Est.) **Interim database transfer** – The objective of this transfer from Aepodia to DNAlytics is to apply the interim analysis plan described in this document. Those patients data are anonymized by Aepodia prior to the transfer to DNAlytics. Those patients will however be identified by their inclusion number.

Chemin du Cyclotron, 6 – B-1348 – Louvain-la-Neuve – Belgium
tva: be 0844.973.235 – iban: be29 5230 8050 8964 – bic: TRIOBEBB
DNAlytics SA is a Spin-Off Company of the Université catholique de Louvain, ICTEAM/INGI