CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

210365Orig1s000

STATISTICAL REVIEW(S)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Science
Office of Biostatistics

Statistical Review and Evaluation

CARCINOGENICITY STUDY

IND/NDA Number:	NDA 210365
Drug Name:	(b) (4) *(Cannabidiol)*
Indication(s):	Adjunctive treatment of seizures associated with Dravet Syndrome or Lennox Gastaut Syndrome.
Studies	Two Year Inhalation Carcinogenicity Study of a Nebulized Aerosol Formulation in the Albino Rat and Mouse.
Applicant:	Sponsor: GW Research LTD c/o GW Pharmaceuticals PLC 68 T.W. Alexander Drive P.O. Box 13628 Research Triangle Park, North Carolina 27709
Review Priority:	Standard
Biometrics Division:	Division of Biometrics VI
Statistical Reviewers:	Steven Thomson Eiji Ishida
Concurring Reviewer:	Karl Lin, Ph.D.
Medical Division:	Division of Neurology Products
Reviewing Pharmacologist:	J. Edward Fisher, PhD
Project Manager:	Stephanie N. Parncutt
Keywords:	Carcinogenicity, Dose response

Table of Contents

1. Background

In this submission, the sponsor included a report and data for a two year carcinogenicity study in rats and a protocol for a similar study in mice. However, the mouse study specifies a 104 week study duration with a target date to issue the protocol in October 2017, indicating the study will not be completed for another year.

The rat study was intended to assess the carcinogenic potential of (b) (4), when administered orally with diet for 104 weeks. The Sponsor notes that one could consider the study as two separate experiments, one in male rats and one in female rats. The Sponsor describes the overall basic structure of the study as: "Two hundred and fifteen male and two hundred and fifteen female rats of the HsdBrlHan:WIST derived strain were allocated to the study and divided into seven groups. Groups 1 to 4 were designated main study animals and Groups 5 to 7 were designated satellite animals and used only for toxicokinetic evaluation." (page 12 of rat report, JJG0003) The study structure is summarized below:

"Dose levels were selected on the basis of a 13-week preliminary study conducted at (b) (4) Study Number JJG0002). The highest dose level of 50 mg/kg/day was expected to provide plasma exposure for parent compound (in terms of AUC0-24) that is significantly (more than 25 times) higher than that anticipated in human use. It was also anticipated that this dose level would produce a 10 % to 20 % deficit in bodyweight gain during the growth phase of males and during the early part of the study for females. The low dose level was chosen as a dose anticipated to produce minimal or no effect in either sex.

"Dose administration began on 14 August 2002 and the necropsies were performed on 11 – 20 August 2004." (page 15 of JJ0003 rat report)

During the administration period, all signs of reaction to treatment were recorded daily. The Sponsor notes that "detailed clinical examinations were recorded weekly and from Week 27 onwards" (page 12 of JJG0003 report). "Blood samples were taken from the satellite animals during Weeks 14, 27, 52 and 105 for assessment of the absorption of the test article"(page 13), as well as vehicle animals in Week 105. Allocation of animals to group was randomized, and animals were housed in groups of 5 by gender.

"The test article was administered in the diet and the formulated diets were freely available for at least 104 weeks, up to the day of necropsy. [Vehicle] Control animals received untreated diet only. Dietary concentrations were adjusted weekly to maintain constant dose levels in relation to bodyweight." (page 20 of report)

"At the end of the treatment period all surviving main study animals were sacrificed by exposure to carbon dioxide gas in a rising concentration and subject to necropsy. A range of tissues was preserved and tissues from all high dose and Control animals were examined microscopically. In addition gross lesions and livers were examined for animals from the low and intermediate dose group and thyroids were examined from low and intermediate dose males." (page 13 of report)

The Sponsor reports that during the administration period, all animals were checked for morbidity, mortality, injury, twice daily, once in the morning and once in the afternoon. Animals were not removed from cage during observation, unless necessary for identification or confirmation of possible findings. The animals were removed from the cage, and detailed observations were conducted for each

animal weekly, beginning during Week 1. The presence of palpable masses was observed during the detailed examination; the site, size and appearance of these masses were recorded when first detected and, following this initial description, the presence or disappearance of these masses were monitored. Any animal showing signs of severe debility or intoxication, and if determined to be moribund or suffering excessively will be euthanized. All animals were subjected to a complete necropsy examination, which included evaluation of the carcass and musculoskeletal system; all external surfaces and orifices; cranial cavity and external surfaces of the brain; and thoracic, abdominal, and pelvic cavities with their associated organs and tissues. Histopathological examinations were performed on all animals found dead, killed moribund, or sacrificed at the end of the experiment, all suspected tumors were diagnosed, and the incidences of benign and malignant tumors of different cell types in the various treatment groups were tabulated. Body weights of individual animals were recorded weekly, for the first 14 weeks, starting during the last week of the prestudy period, and then monthly thereafter, as well as on the day of necropsy. Terminal body weights were not collected from animals found dead or euthanized moribund.

2. Sponsor's Analyses

2.1. Sponsor's Survival analyses

The "Sponsor's analysis showed the numbers of rats surviving to their terminal necropsy were 26 (43%), 33 (55%), 29 (48%), and 28 (47%) in the reference control group, low, medium, and high dose groups, in male rats, respectively, and 21 (35%), 29 (48%), 31 (52%) and 24 (40%) in reference control, low, medium, and high dose groups, in female rats, respectively. The sponsor's report showed no significance at the 5% level using a log-rank test, for both male and female datasets, with p-value = 0.7020 and p-value = 0.3616 respectively. Therefore, no post-hoc testing was done for these datasets, i.e. neither the trend test nor the pairwise comparisons [for survival differences] were performed." (page 22 of Sponsor's report)

2.2. Sponsor's Tumor data analysis

Statistical analyses of carcinogenicity include tests for dose-response relationship (positive trend) among the increasing doses over time and pairwise comparisons of treated groups with control in tumor incidence by organ/tumor combination. There are two major concerns in analyzing such carcinogenicity data, namely, adjustment for the difference in mortality due to drug toxicity and adjustment for the multiplicity due to multiple testing of trends and pairwise differences by organ tumor combination. The Sponsor's report states that the "analysis of the tumour incidence data, for both neoplastic and non-neoplastic lesions, was performed in line with the methodology given in the IARC annex (Peto et al, 1980). The Peto test adjusts the mortality differences among treatment groups by partitioning the entire study period into several intervals, analyzing the data separately for each interval, and then combining them using the Mantel-Haenszel procedure. Statistical analyses of include tests for dose-response relationship (positive trend) among the increasing doses over time and pairwise comparisons of treated groups with control in tumor incidence by organ/tumor combination. The denominator for the calculation of the proportion of tumor bearing animals is determined from the cause of death information tumor data.

There has been concern regarding the construction of suitable intervals for mortality adjustment. Further it seems to be difficult to accurately specify retrospectively when a tumor is the real cause of death of an animal. Thus this information may be quite imprecise. For this reason it seems that the results of carcinogenicity analyses using the Peto test have been questioned due to the likelihood of inaccurate cause of death information.

The Sponsor summarizes their carcinogenicity results as follows: "There were no increases in findings that were considered to be indicative of carcinogenic potential when the neoplastic findings recorded in treated animals were compared with Controls. However, the administration of the test article appeared to have had a beneficial effect in respect of commonly encountered tumours that are associated with hormonally mediated neoplasia in ageing rats.

"The incidence of pituitary tumours was lower in both sexes; relatively more so in males than females. There was however a higher incidence of pituitary hyperplasia in the treated males, perhaps indicating that the beneficial effect of treatment is a delay in the onset of such proliferative lesions rather than a complete inhibition.

"Overall, the number of mammary tumours in high dose females was about a third of those seen in Control females. There was a general reduction in activity of mammary tissue in both sexes (acinar proliferation/secretory activity), although this was equivocal in males. Prolactin driven hyperactivity is widely recognised as a precursor to mammary neoplasia, suggesting these changes may be the result of small changes to the hormonal status in treated animals, although a similar effect may be seen as a result of reduced bodyweight gain.

"There was a higher incidence of benign granulosa cell tumours in the ovaries of the high dose females compared to Controls (3/50:0/50). This tumour appears to occur in clusters occasionally, and in a recent study conducted at this laboratory was seen in 4/46 control females. Given that there was no increase in proliferative lesions in this organ this higher incidence is considered to represent a similar cluster. "There were no unusual types of neoplasia seen after oral administration of [the test drug] … for 104 weeks and the incidence and distribution of the other tumours recorded were those expected in this strain of rat in this laboratory." (page 36 of JJG0003 of report).

3. FDA Reviewer's analyses

To verify sponsor's analysis, the FDA reviewers independently performed the survival and tumor data analyses.

3.1 FDA Survival analysis

Again, the numbers of rats surviving to their terminal necropsy were 26 (43%), 33 (55%), 29 (48%), and 28 (47%) in the reference vehicle control group, low, medium, and high dose groups, in male rats, respectively, and 21 (35%), 29 (48%), 31 (52%) and 24 (40%) in reference control, low, medium, and high dose groups, in female rats, respectively. This reviewer's analysis showed no statistically significant increase or decrease in mortality across the reference control group and the three treated groups in either sex of rats. The pairwise comparisons showed no statistically significant increase or decrease in mortality between each of the treated groups and the reference control group in either sex of rats.

3.2 FDA Tumor Data Analysis

An alternative to the Peto test was suggested by Bailer and Portier, popularly known as the poly-k test and is described in some detail below. Unlike the Peto test, this test does not need any arbitrary partitioning of the study period or the cause of death information. The poly-k test for trend in tumor incidence adjusts the differences in mortality among treatment groups by assigning a weight of less than one to an animal that died early without developing the tumor; and a weight of one to an animal that died with the tumor or survived to the end of the study. This is described in some detail below:

The test for trend in tumor incidence adjusts the differences in mortality among treatment groups by assigning a weight of less than one to an animal that died early without developing the tumor; and a weight of one to an animal that died with the tumor or survived to the end of the study. The sum of the assigned weights of animals in a treatment group is then used as the denominator for the calculation of proportion of tumor-bearing animals for the group. The less-than-one weight assigned to an animal is the fraction of the animal's surviving time in the study over the maximum time of the study with a power k. The power k of the fraction is determined by the distribution of tumor onset times of the tumor. The Poly-k test seems to have advantages over the Peto test in the sense that it does not require the cause of death information, which is an essential part for the Peto test. Drs. Lin and Rahman at the FDA compared the overall false positive rates of the Peto and Poly-K tests using the Rahman-Lin multiple comparison adjustment based on some simulation results.

Unlike the Peto test, this test does not need any arbitrary partitioning of the study period or the cause of death information. In this method, an animal that lives the full study period ($w_{\max}$) or dies before the terminal sacrifice with development of the tumor type being tested gets a score of s_h =1. An animal that dies at Week w_h without development of the given tumor type before the end of the study gets a score of s_h $=\left(\frac{w_h}{w_{\max}}\right)^k$ <1. The adjusted group size is defined as Σs_h . As an interpretation, an animal with score s_h =1 can be considered as a whole animal, while an animal with score s_h <1 can be considered as a part of an animal. The adjusted group size Σs_h is equal to N (the original group size) if all animals live up to the end of the study or if each animal develops the given tumor being tested, otherwise the adjusted group size is less than N. These adjusted group sizes are then used for the dose response relationship (or the pairwise comparison) tests using the Cochran-Armitage test. One critical point for Poly-k test is the choice of the appropriate value of k. For long term 104-week standard rat and mouse studies, a value of k=3 is suggested in the literature [Gebregziabher and Hoel (2009), Moon et al. (2003), Portier, et al. (1986)]. Hence, this reviewer used k=3 for the analysis of the data. Based on the intent to treat (ITT) principle w_{max} was considered as 105 for both male and female rats.

The tumor rates and the p-values of the tested tumor types are listed in Tables 3A and 3B in Appendix 1 for female rats and male rats, respectively.

Multiple testing adjustments:

Following the FDA revised draft guidance (Lin) for the carcinogenicity study design and data analysis, for the standard two-year rodent study significance levels of 0.005 and 0.025 for common and rare tumors, respectively in dose response relationship (trend) tests and significance levels of 0.01 and 0.05 for common and rare tumors, respectively in pairwise comparisons. A tumor is defined as a rare tumor if the published spontaneous rate or the spontaneous rate of the reference control of the tumor is less than 1%, and a common tumor is defined as one with tumor rate greater than or equal to 1%.

FDA Reviewer's findings:

The tumor types with p-values less than 0.05 for dose response relationship and/or pairwise comparisons of reference control and treated groups are reported in Table 2, below:

Table 2: Organ-Tumor Types with Statistically Significant tests for Dose Trend or the Pairwise Comparisons

Treated Groups and Reference control Group in Rats

Gender	Organ Name	Tumor Name	0 µg/kg Veh. Cont. (N=50) P - Trend	30 µg/kg Low (N=50) P - RC vs. L	100 µg/kg Med (N=50) P - RC vs. M	300 µg/kg High (N=50) P - RC vs. H
Male	Hemopoietic Tissue	Large Granular Lymphocyte Leukemia Malignant Tumor	0/50 (44) 0.4817	2/16 (11) 0.0370	0/9 (5)	1/50 (46) 0.5172
Female	Ovaries	Tubulostromal Adenoma Benign Tumor	0/50 (45) 0.0425	1/50 (42) 0.4828	1/50 (47) 0.5109	3/50 (44) 0.1166
	Thyroid	Follicular Adenoma Benign Tumor	0/50 (45) 0.0188	0/50 (42)	1/50 (47) 0.5109	3/50 (44) 0.1166

X/ZZ (YY): X=number of tumor bearing animals; YY=mortality weighted total number of animals; ZZ=unweighted total number of animals observed;

Following the multiple testing adjustment method described above, the only statistically significant test was the test of benign follicular adenoma in the thyroid of females classified as a rare tumor, (0.0188 < 0.025). This analysis showed no other tumor types with either a statistically significant trend in tumor incidence in response to dose or a statistically significant pairwise comparison of each dose to the vehicle control.

Appendix 1
Table 3.A Carcinogenicity in Female Rats

Organ Name Organ Name	Tumor Name Tumor Name	0 mg Cont (N=60) P - Trend Control	30 mg Low (N=60) P - C vs. L Low Dose	100 mg Med (N=60) P - C vs. M Mid Dose	300 mg High (N=60) P - C vs. H High Dose
Abdominal cavity	LIPOMA - BENIGN TUMOUR	0/50 (45) 0.6283	1/50 (42) 0.4828	1/50 (47) 0.5109	0/50 (44)
Adrenals	CORTICAL ADENOMA - BENIGN TUMOUR	2/50 (45) 0.6714	1/50 (42) 0.8661	0/50 (47) 1.0000	1/50 (44) 0.8750
Brain	ASTROCYTOMA - MALIGNANT TUMOUR	0/50 (45) 0.2472	0/50 (42)	0/50 (47)	1/50 (44) 0.4944
	OLIGODENDROGLIOMA - MALIGNANT TUMOUR	0/50 (45) 0.5112	0/50 (42)	1/50 (47) 0.5109	0/50 (44)
Cervix	HISTIOCYTIC SARCOMA - MALIGNANT TUMOUR	0/50 (45) 0.5112	0/50 (42)	1/50 (47) 0.5109	0/50 (44)
	LEIOMYOSARCOMA - MALIGNANT TUMOUR	0/50 (45) 0.5112	0/50 (42)	1/50 (47) 0.5109	0/50 (44)
	POLYP - BENIGN TUMOUR	0/50 (45) 0.5112	0/50 (42)	1/50 (47) 0.5109	0/50 (44)
Clitoral glands	ADENOCARCINOMA - MALIGNANT TUMOUR	0/50 (45) 0.2486	0/49 (41)	0/50 (47)	1/50 (44) 0.4944
	DUCTAL ADENOMA - BENIGN TUMOUR	0/50 (45) 0.2486	0/49 (41)	0/50 (47)	1/50 (44) 0.4944
	PAPILLOMA - BENIGN TUMOUR	1/50 (45) 1.0000	0/49 (41) 1.0000	0/50 (47) 1.0000	0/50 (44) 1.0000
Duodenum	ADENOCARCINOMA - MALIGNANT TUMOUR	0/50 (45) 0.2472	0/50 (42)	0/50 (47)	1/50 (44) 0.4944
Femur & joint (incl. marrow)	HAEMANGIOMA - BENIGN TUMOUR	1/50 (45) 1.0000	0/50 (42) 1.0000	0/50 (47) 1.0000	0/50 (44) 1.0000
Liver	ADENOMA - BENIGN TUMOUR	0/50 (45) 0.5112	0/50 (42)	1/50 (47) 0.5109	0/50 (44)
	HAEMANGIOSARCOMA - MALIGNANT TUMOUR	0/50 (45) 0.2472	0/50 (42)	0/50 (47)	1/50 (44) 0.4944
Mesenteric lymph nodes	HAEMANGIOMA - BENIGN TUMOUR	3/50 (45) 1.0000	0/50 (42) 1.0000	0/50 (47) 1.0000	0/50 (44) 1.0000
Nasal cavity	OLFACTORY NEUROBLASTOMA - MALIGNANT TUMOUR	0/50 (45) 0.5112	0/50 (42)	1/50 (47) 0.5109	0/50 (44)

Organ Name Organ Name	Tumor Name Tumor Name	0 mg Cont (N=60) P - Trend Control	30 mg Low (N=60) P - C vs. L Low Dose	100 mg Med (N=60) P - C vs. M Mid Dose	300 mg High (N=60) P - C vs. H High Dose
Ovaries	GRANULOSAR THECAL CELL TUMOUR - BENIGN TUMOUR	0/50 (45) 0.2472	0/50 (42)	0/50 (47)	1/50 (44) 0.4944
	TUBULOSTROMAL ADENOMA - BENIGN TUMOUR	0/50 (45) 0.0425	1/50 (42) 0.4828	1/50 (47) 0.5109	3/50 (44) 0.1166
Pancreas	ISLET CELL ADENOMA - BENIGN TUMOUR	1/50 (45) 0.9372	1/50 (42) 0.7354	0/50 (47) 1.0000	0/50 (44) 1.0000
	ISLET CELL CARCINOMA - MALIGNANT TUMOUR	0/50 (45) 0.7472	1/50 (42) 0.4828	0/50 (47)	0/50 (44)
Pituitary gland	ADENOCARCINOMA - MALIGNANT TUMOUR	2/50 (45) 0.7742	6/50 (43) 0.1187	0/50 (47) 1.0000	2/49 (44) 0.6833
	ADENOMA - BENIGN TUMOUR	29/50 (49) 0.9293	27/50 (47) 0.6478	20/50 (48) 0.9734	20/49 (45) 0.9493
Site of Mammary Gland	MAMMARY FIBROADENOMA - BENIGN TUMOUR	0/50 (45) 0.3086	1/50 (42) 0.4828	0/50 (47)	1/50 (44) 0.4944
Stomach	FIBROSARCOMA - MALIGNANT TUMOUR	1/50 (46) 1.0000	0/50 (42) 1.0000	0/50 (47) 1.0000	0/50 (44) 1.0000
Subcutaneous fat	LIPOMA - BENIGN TUMOUR	0/50 (45) 0.5112	0/50 (42)	1/50 (47) 0.5109	0/50 (44)
	MAMMARY FIBROADENOMA - BENIGN TUMOUR	5/50 (45) 0.9979	3/50 (42) 0.8437	1/50 (47) 0.9886	0/50 (44) 1.0000
Thymus	THYMOMA - BENIGN TUMOUR	1/50 (45) 0.3496	2/50 (42) 0.4738	1/50 (47) 0.7635	2/50 (44) 0.4915
	THYMOMA - MALIGNANT TUMOUR	0/50 (45) 0.2472	0/50 (42)	0/50 (47)	1/50 (44) 0.4944
Thyroid glands	C CELL ADENOMA - BENIGN TUMOUR	2/50 (45) 0.2747	2/50 (42) 0.6653	1/50 (47) 0.8870	3/50 (44) 0.4892
	FOLLICULAR ADENOMA - BENIGN TUMOUR	0/50 (45) 0.0188	0/50 (42)	1/50 (47) 0.5109	3/50 (44) 0.1166
Uterus	ADENOCARCINOMA - MALIGNANT TUMOUR	1/50 (45) 0.1682	0/50 (42) 1.0000	1/50 (47) 0.7635	2/50 (44) 0.4915
	ADENOMA - BENIGN TUMOUR	2/50 (45) 1.0000	0/50 (42) 1.0000	0/50 (47) 1.0000	0/50 (44) 1.0000
	LEIOMYOSARCOMA - MALIGNANT TUMOUR	0/50 (45) 0.5112	0/50 (42)	1/50 (47) 0.5109	0/50 (44)

Organ Name Organ Name	Tumor Name Tumor Name	0 mg Cont (N=60) P - Trend Control	30 mg Low (N=60) P - C vs. L Low Dose	100 mg Med (N=60) P - C vs. M Mid Dose	300 mg High (N=60) P - C vs. H High Dose
	POLYP - BENIGN TUMOUR	2/50 (45) 0.1876	2/50 (42) 0.6653	4/50 (47) 0.3595	4/50 (44) 0.3275
Vagina	POLYP - BENIGN TUMOUR	1/50 (46) 1.0000	0/50 (42) 1.0000	0/50 (47) 1.0000	0/50 (44) 1.0000
non-protocol skin	ANAPLASTIC CARCINOMA - MALIGNANT TUMOUR	0/50 (45) 0.2472	0/50 (42)	0/50 (47)	1/50 (44) 0.4944
	BASAL CELL TUMOUR - MALIGNANT TUMOUR	1/50 (45) 1.0000	0/50 (42) 1.0000	0/50 (47) 1.0000	0/50 (44) 1.0000
	FIBROSARCOMA - MALIGNANT TUMOUR	0/50 (45) 0.7472	1/50 (42) 0.4828	0/50 (47)	0/50 (44)
	KERATOACANTHOMA - BENIGN TUMOUR	1/50 (45) 0.7625	0/50 (42) 1.0000	1/50 (47) 0.7635	0/50 (44) 1.0000
	MAMMARY ADENOCARCINOMA - MALIGNANT TUMOUR	3/50 (45) 0.8822	3/50 (42) 0.6282	1/50 (47) 0.9467	1/50 (44) 0.9390
	MAMMARY ADENOMA - BENIGN TUMOUR	2/50 (45) 0.1611	1/50 (42) 0.8661	0/50 (47) 1.0000	3/50 (44) 0.4892
	MAMMARY FIBROADENOMA - BENIGN TUMOUR	15/50 (46) 0.9915	16/50 (44) 0.4391	13/50 (47) 0.7721	7/50 (46) 0.9868
	PAPILLOMA - BENIGN TUMOUR	1/50 (45) 1.0000	0/50 (42) 1.0000	0/50 (47) 1.0000	0/50 (44) 1.0000
	SQUAMOUS CELL CARCINOMA - MALIGNANT TUMOUR	0/50 (45) 0.7472	1/50 (42) 0.4828	0/50 (47)	0/50 (44)

Table 3.B Carcinogenicity in Male Rats

Organ Name	Tumor Name	Control	Low Dose	Mid Dose	High Dose
Abdominal fat	FIBROSARCOMA - MALIGNANT TUMOUR	0/50 (44) 0.7556	1/50 (44) 0.5000	0/50 (46)	0/50 (46)
Adrenals	ADENOCARCINOMA - MALIGNANT TUMOUR	0/50 (44) 0.5111	0/50 (44)	1/50 (46) 0.5111	0/50 (46)
	CORTICAL ADENOMA - BENIGN TUMOUR	1/50 (45) 0.2439	1/50 (44) 0.7472	0/50 (46) 1.0000	2/50 (46) 0.5083
	PHAEOCHROMOCYTOMA - BENIGN TUMOUR	2/50 (44) 0.5898	0/50 (44) 1.0000	0/50 (46) 1.0000	1/50 (46) 0.8873
	PHAEOCHROMOCYTOMA - MALIGNANT TUMOUR	1/50 (44) 0.3668	2/50 (44) 0.5000	1/50 (46) 0.7638	2/50 (46) 0.5169
Brain	GRANULAR CELL TUMOUR - MALIGNANT TUMOUR	0/50 (44) 0.2556	0/50 (44)	0/50 (46)	1/50 (46) 0.5111
Caecum	CARCINOMA - MALIGNANT TUMOUR	0/50 (44) 0.7569	1/50 (45) 0.5056	0/50 (46)	0/50 (46)
Diaphragm	MESOTHELIOMA - MALIGNANT TUMOUR	0/50 (44) 0.2556	0/50 (44)	0/50 (46)	1/50 (46) 0.5111
Epididymal fat	MESOTHELIOMA - MALIGNANT TUMOUR	0/50 (44) 0.5111	0/50 (44)	1/50 (46) 0.5111	0/50 (46)
	SCHWANNOMA - MALIGNANT TUMOUR	0/50 (44) 0.7556	1/50 (44) 0.5000	0/50 (46)	0/50 (46)
Epididymides	MESOTHELIOMA - MALIGNANT TUMOUR	0/50 (44) 0.2556	0/50 (44)	0/50 (46)	1/50 (46) 0.5111
Heart	ENDOCARDIAL SCHWANNOMA - MALIGNANT TUMOUR	0/50 (44) 0.5111	0/50 (44)	1/50 (46) 0.5111	0/50 (46)
Hindlimbs	HAEMANGIOMA - BENIGN TUMOUR	0/50 (44) 0.7556	1/50 (44) 0.5000	0/50 (46)	0/50 (46)
Kidneys	LIPOSARCOMA - MALIGNANT TUMOUR	1/50 (44) 1.0000	0/50 (44) 1.0000	0/50 (46) 1.0000	0/50 (46) 1.0000
	RENAL MESENCHYMAL TUMOUR - MALIGNANT TUMOUR	1/50 (44) 1.0000	0/50 (44) 1.0000	0/50 (46) 1.0000	0/50 (46) 1.0000

Organ Name	Tumor Name	Control	Low Dose	Mid Dose	High Dose
Liver	CHOLANGIOCARCINOMA - MALIGNANT TUMOUR	1/50 (45) 1.0000	0/50 (44) 1.0000	0/50 (46) 1.0000	0/50 (46) 1.0000
Mesenteric lymph nodes	HAEMANGIOMA - BENIGN TUMOUR	3/50 (44) 0.8006	1/50 (44) 0.9418	1/50 (46) 0.9469	1/50 (46) 0.9469
Nasal cavity	OLFACTORY NEUROBLASTOMA - MALIGNANT TUMOUR	0/50 (44) 0.5111	0/50 (44)	1/50 (46) 0.5111	0/50 (46)
	SCHWANNOMA - MALIGNANT TUMOUR	1/50 (45) 1.0000	0/50 (44) 1.0000	0/50 (46) 1.0000	0/50 (46) 1.0000
Oral cavity	SQUAMOUS CELL CARCINOMA - MALIGNANT TUMOUR	0/50 (44) 0.2597	0/50 (44)	0/50 (46)	1/50 (47) 0.5165
Pancreas	EXOCRINE CELL ADENOMA - BENIGN TUMOUR	1/50 (44) 1.0000	0/50 (44) 1.0000	0/50 (46) 1.0000	0/50 (46) 1.0000
	ISLET CELL ADENOMA - BENIGN TUMOUR	1/50 (45) 0.9605	3/50 (44) 0.2997	0/50 (46) 1.0000	0/50 (46) 1.0000
	ISLET CELL CARCINOMA - MALIGNANT TUMOUR	1/50 (44) 1.0000	0/50 (44) 1.0000	0/50 (46) 1.0000	0/50 (46) 1.0000
Pituitary gland	ADENOMA - BENIGN TUMOUR	14/47 (45) 0.9281	8/50 (44) 0.9522	8/50 (47) 0.9668	7/50 (47) 0.9828
Preputial glands	ADENOMA - BENIGN TUMOUR	1/50 (44) 1.0000	0/50 (44) 1.0000	0/50 (46) 1.0000	0/50 (46) 1.0000
	CARCINOMA - MALIGNANT TUMOUR	0/50 (44) 0.3212	1/50 (44) 0.5000	0/50 (46)	1/50 (46) 0.5111
Prostate gland	ADENOMA - BENIGN TUMOUR	0/50 (44) 0.2556	0/50 (44)	0/50 (46)	1/50 (46) 0.5111
Spleen	HAEMANGIOMA - BENIGN TUMOUR	1/50 (44) 1.0000	0/50 (44) 1.0000	0/50 (46) 1.0000	0/50 (46) 1.0000
	HAEMANGIOSARCOMA - MALIGNANT TUMOUR	2/50 (44) 1.0000	0/50 (44) 1.0000	0/50 (46) 1.0000	0/50 (46) 1.0000
Subcutaneous fat	HIBERNOMA - BENIGN TUMOUR	0/50 (44) 0.5111	0/50 (44)	1/50 (46) 0.5111	0/50 (46)
	SCHWANNOMA - MALIGNANT TUMOUR	1/50 (44) 1.0000	0/50 (44) 1.0000	0/50 (46) 1.0000	0/50 (46) 1.0000
Tail	KERATOACANTHOMA - BENIGN TUMOUR	3/49 (43) 1.0000	0/50 (44) 1.0000	0/50 (46) 1.0000	0/50 (46) 1.0000
Testes	INTERSTITIAL CELL TUMOUR - BENIGN TUMOUR	4/50 (44) 0.8701	1/50 (44) 0.9723	0/50 (46) 1.0000	1/50 (46) 0.9753

Organ Name	Tumor Name	Control	Low Dose	Mid Dose	High Dose
Thoracic cavity	LIPOMA - BENIGN TUMOUR	0/50 (44) 0.5111	0/50 (44)	1/50 (46) 0.5111	0/50 (46)
Thymus	THYMOMA - BENIGN TUMOUR	1/50 (45) 0.5245	1/50 (44) 0.7472	1/49 (45) 0.7528	1/50 (46) 0.7582
	THYMOMA - MALIGNANT TUMOUR	0/50 (44) 0.5111	0/50 (44)	1/49 (46) 0.5111	0/50 (46)
Thyroid glands	C CELL ADENOMA - BENIGN TUMOUR	0/50 (44) 0.3197	3/50 (44) 0.1207	1/50 (46) 0.5111	2/50 (46) 0.2584
	FOLLICULAR ADENOMA - BENIGN TUMOUR	6/50 (45) 0.1952	6/50 (44) 0.6050	8/50 (46) 0.4036	9/50 (46) 0.3028
	FOLLICULAR CARCINOMA - MALIGNANT TUMOUR	0/50 (44) 0.7556	1/50 (44) 0.5000	0/50 (46)	0/50 (46)
Zymbal's Gland	CARCINOMA - MALIGNANT TUMOUR	0/48 (42) 0.3236	1/50 (44) 0.5116	0/50 (46)	1/49 (45) 0.5172
haemopoietic tissue	LARGE GRANULAR LYMPHOCYTE LEUKAEMIA - MALIGNANT TUMOUR	0/50 (44) 0.4817	2/16 (11) 0.0370	0/9 (5)	1/50 (46) 0.5111
	LYMPHOMA - MALIGNANT TUMOUR	0/50 (44) 0.8235	1/16 (10) 0.1852	2/9 (6) 0.0122	0/50 (46)
non-protocol skin	BASAL CELL TUMOUR - BENIGN TUMOUR	1/50 (44) 1.0000	0/50 (44) 1.0000	0/50 (46) 1.0000	0/50 (46) 1.0000
	FIBROMA - BENIGN TUMOUR	2/50 (44) 0.6725	1/50 (44) 0.8793	2/50 (47) 0.7163	1/50 (46) 0.8873
	FIBROSARCOMA - MALIGNANT TUMOUR	0/50 (44) 0.5111	0/50 (44)	2/50 (47) 0.2640	0/50 (46)
	HISTIOCYTIC SARCOMA - MALIGNANT TUMOUR	0/50 (44) 0.2556	0/50 (44)	0/50 (46)	1/50 (46) 0.5111
	KERATOACANTHOMA - BENIGN TUMOUR	6/50 (44) 0.6318	3/50 (44) 0.9217	3/50 (47) 0.9356	4/50 (46) 0.8601
	LIPOMA - BENIGN TUMOUR	1/50 (44) 0.9426	2/50 (44) 0.5000	0/50 (46) 1.0000	0/50 (46) 1.0000
	LIPOSARCOMA - MALIGNANT TUMOUR	0/50 (44) 0.7556	1/50 (44) 0.5000	0/50 (46)	0/50 (46)
	OSTEOSARCOMA - MALIGNANT TUMOUR	1/50 (45) 0.9399	1/50 (45) 0.7528	0/50 (46) 1.0000	0/50 (46) 1.0000
	PAPILLOMA - BENIGN TUMOUR	1/50 (44) 0.1610	0/50 (44) 1.0000	0/50 (46) 1.0000	2/50 (46) 0.5169

Organ Name	Tumor Name	Control	Low Dose	Mid Dose	High Dose
	SQUAMOUS CELL CARCINOMA - MALIGNANT TUMOUR	0/50 (44) 0.7569	1/50 (45) 0.5056	0/50 (46)	0/50 (46)

Appendix 2. References

1. Bailer, A. and Portier, C. (1988), "Effects of Treatment-Induced Mortality on Tests for Carcinogenicity in Small Samples", *Biometrics*, **44**, 4, 417-431.

2. Bieler, G.S., and Williams, R.L. (1993), "Ratio Estimates, the Delta Method, and Quantal Response Tests for Increased Carcinogenicity", *Biometrics*, **49**, 4, 793-801.

3. Chu, K.C., Ceuto, C., and Ward, J.M. (1981), "Factors in the Evaluation of 200 National Cancer Institute Carcinogen Bioassays", *Journal of Toxicology and Environmental Health,* **8**, 51-280.

4. Haseman, J. K. (1983), "A Reexamination of False-positive Rates for Carcinogenicity Studies", *Fundamental and Applied Toxicology,* **3**, 334-339.

5. Lin, K. K. and Ali, M.W. (2006), "Statistical Review and Evaluation of Animal Tumorigenicity Studies", *Statistics in the Pharmaceutical Industry, Third Edition,* edited by C.R. Buncher and J.Y. Tsay, Marcel Dekker, Inc. New York.

6. Lin, K. K. and Rahman, M.A. (1998), "Overall False Positive Rates in Tests for Linear Trend in Tumor Incidence in Animal Carcinogenicity Studies of New Drugs", *Journal of Biopharmaceutical Statistics*. **8**, 1, 1-15.

7. Peto, R., Pike, M.C., Day, N.E., Gray, R.G., Lee, P.N., Parrish, S., Peto, J., Richards, S., and Wahrendorf, J. (1980). "Guidelines for sample sensitive significance tests for carcinogenic effects in long-term animal experiments", *IARC Monographs on the Evaluation of Carcinogenic Risk of Chemicals to Humans, supplement 2: Long term and Short term Screening Assays for Carcinogens: A Critical Appraisal*, International Agency for Research Against Cancer, 311-426.

8. Rahman, M.A. and Lin, K.K. (2008), "A Comparison of False Positive Rates of Peto and Poly-3 Methods for Long Term Carcinogenicity Data Analysis Using Multiple Comparison Adjustment Method Suggested by Lin and Rahman", *Journal of Biopharmaceutical Statistics*. **18**, 949-958.

9 STP Peto Working Group (2002), "Statistical Methods for Carcinogenicity Studies", *Toxicologic Pathology*. **30** (3), 403-414.

10. U.S. Department of Health and Human Services (2013), Guidance for Industry Statistical Aspects of the Design, Analysis, and Interpretation of Chronic Rodent Carcinogenicity Studies of Pharmaceuticals (DRAFT GUIDANCE), Center for Drug Evaluation and Research, Food and Drug Administration.

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

STEVEN F THOMSON
06/25/2018

KARL K LIN
06/26/2018
Concur with review.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Science
Office of Biostatistics

Statistical Review and Evaluation

CLINICAL STUDIES

NDA/Serial Number:	210365
Drug Name:	GWP42003-P
Indication:	NA
Study number:	GWEP1431
Applicant:	GW Research Ltd
Date(s):	Date of Document: Aug 4, 2017 Consult received date: Nov 2, 2017 Completion date: 03/22/2018
Review Priority:	Priority
Biometrics Division:	DBVI
Statistical Reviewer:	Anna Sun, Ph.D., Mathematical Statistician, OB/DBVI
Concurring Reviewers:	Qianyu Dang, Ph.D., Lead Mathematical Statistician, OB/DBVI Yi Tsong, Ph.D., Division Director, OB/DBVI
Medical Division:	Control Substance Staff
The CSS Team:	Katherine Bonson, Ph.D., Pharmacologist, CDER/OCD/CSS
Project Manager:	Sandra Saltz, Project Manager, CSS

Keywords: *Crossover design, Drug abuse potential study, Self-reported endpoint, Multiple endpoints*

Table of Contents

List of Tables

List of Figures

1. Executive Summary

Study GWEP1431 was a single-site, single-dose, randomized, double-blind, double-dummy, placebo- and active-controlled crossover trial. The abuse potential of single oral doses of GWP42003-P (750, 1500, and 4500 mg) were compared with that of single oral doses of alprazolam 2 mg, dronabinol 10 and 30 mg, and placebo in healthy recreational polydrug users. Subjects participated in an outpatient medical Screening visit, a 7-day Qualification (Drug Discrimination) Phase, a 7-period Treatment Phase, and an outpatient safety Follow-up visit.

The primary objective of this trial was to evaluate the abuse potential of single doses of GWP42003-P compared with alprazolam, dronabinol and placebo in healthy recreational polydrug users.

The Emax on the bipolar Drug Liking VAS was the primary pharmacodynamic endpoint.

The treatment comparisons to assess the abuse potential of GWP42003-P included the following:

• Alprazolam vs. placebo (trial validity)

• Each dose of dronabinol vs. placebo (trial validity)

• Each dose of GWP42003-P vs. placebo (absolute abuse potential)

• Each dose of GWP42003-P vs. alprazolam (relative abuse potential)

• Each dose of GWP42003-P vs. each dose of dronabinol (relative abuse potential)

The reviewer analyzed the primary PD endpoint Drug Liking, and the secondary PD endpoints: High, Good Effects, Take Drug Again and Overall Drug Liking. The results from the statistical reviewer's analyses establish that:

- The study met the validity criteria based on the statistically significant differences in Drug Liking Emax between alprazolam 2 mg versus placebo and dronabinol 30 mg versus placebo. The lowest dose of the positive control, dronabinol 10 mg, did not meet the study validity criteria using a 15-point margin for Drug Liking VAS Emax compared with placebo. Since the study was designed and conducted based on the recommendations in the draft guidance on the Assessment of Abuse Potential of Drugs (Jan 2010) and completed before the final guidance was issued in Jan 2017. Thus, the study was not powered with an adequate sample size to perform this post-hoc analysis, which should be considered in the interpretation of this result.

- All 3 GWP42003-P doses were associated with significantly lower effects than the positive controls on the primary endpoint and secondary endpoints of High, Good Effects, Take Drug Again and Overall Drug Liking (P value <0.01), with the exception of dronabinol 10 mg versus GWP42003-P 4500 mg for High (P value=0.1544).

- Overall, GWP42003-P 750 mg showed little significant and no consistent abuse potential. However, higher doses of GWP42003-P (1500 and 4500 mg) demonstrated significantly greater effects compared with placebo on the primary endpoint and secondary endpoints of High, Good Effects and Take Drug Again. **Higher doses of GWP42003-P (1500 and 4500 mg) are associated with a signal for abuse potential.**

2. Review Report on Study GWEP1431

2.1 Introduction

The cannabis plant (Cannabis sativa L.) produces trichomes that synthesize a large number of pharmacologically active compounds called phytocannabinoids. The most abundant of these are Δ^9-tetrahydrocannabinol (THC) and cannabidiol (CBD), although the amounts and proportions of the various phytocannabinoids in each plant vary by strain and can be adjusted by breeding.

GWP42003-P is formulated from extracts prepared from Cannabis sativa L. plants that have a defined chemical profile and contain consistent levels of CBD as the principal phytocannabinoid. (b) (4) (b) (4) . Note that the IMP will be referred to as GWP42003-P throughout this document, although the source tables, listings and figures present the product as "CBD".

The efficacy and safety of open label GWP42003-P (up to 25 mg/kg/day) was recently reported in children and young adults with drug resistant epilepsy in an expanded access compassionate use program. Safety data from 313 patients showed that GWP42003 P was generally well tolerated at doses up to 25 mg/kg/day with only 4% of patients discontinuing due to an adverse event (AE). Dependent on the trial site, titration to a maximum dose of 50 mg/kg/day was allowed. Based on the available safety data, no dose-related changes in benefit-risk have been established. In a recently completed single ascending dose (SAD) and multiple-ascending dose

(MAD) study of GWP42003-P in healthy subjects (GWEP1544), single doses of GWP42003 P up to 6000 mg were well tolerated, as were multiple doses of 1500 mg twice daily for 6.5 days. There were no serious adverse events (SAEs) or discontinuations due to an AE.

Thus, while there was no evidence to suggest that GWP42003-P is likely to produce euphoria or any other effect associated with abuse or dependence, CBD has shown anxiolytic effects and some reports of drowsiness/somnolence in the published literature. Treatment-emergent AEs of somnolence were also reported with GWP42003-P in the recently completed SAD and MAD trial. Due to the fact that GWP42003-P is a CNS-active drug whose abuse-related subjective effects had not yet been directly evaluated in a controlled clinical trial, the purpose of this trial was to evaluate the abuse potential of single doses of GWP42003-P in healthy recreational polydrug users.

2.1.1 Objectives of the study

The primary objective of this trial was to evaluate the abuse potential of single doses of GWP42003-P compared with alprazolam, dronabinol and placebo in healthy recreational polydrug users.

The secondary objectives of the trial were:

• To evaluate the safety and tolerability of GWP42003-P in healthy recreational polydrug users.

• To evaluate the pharmacokinetics of CBD, THC, and their major metabolites in healthy recreational polydrug users.

2.1.2 Study design

The design was a single-site, single-dose, randomized, double-blind, double-dummy, placebo- and active-controlled crossover trial. The abuse potential of single oral doses of GWP42003-P (750, 1500, and 4500 mg) were compared with that of single oral doses of alprazolam 2 mg, dronabinol 10 and 30 mg, and placebo in healthy recreational polydrug users. Subjects participated in an outpatient medical Screening visit, a 7-day Qualification (Drug Discrimination) Phase, a 7-period Treatment Phase, and an outpatient safety Follow-up visit.

Within 28 days of the Screening visit, eligible subjects were admitted to the clinical research unit (CRU) (Day-1) for the Qualification Phase. During the Qualification Phase, subjects received single oral doses of alprazolam 2 mg, dronabinol 20 mg, and matching placebo in a randomized, double blind, crossover manner, with each drug administration separated by approximately 48 hours (Day 1, Day 3 and Day 5), to ensure that they could discriminate and show positive subjective effects of the active controls. Subjects were discharged from the CRU at approximately 24 hours after the last drug administration (Day 6). A washout interval of at least 8 days (maximum of 21 days) was required between last drug administration in the Qualification Phase and first drug administration in the Treatment Phase.

Following confirmation of eligibility from the Qualification Phase, subjects were randomized to 1 of 14 treatment sequences according to two 7 × 7 Williams squares. Subjects were admitted to the CRU on the day (Day -1) prior to drug administration in each Treatment Period and remained resident until approximately 24 hours after each drug administration (Day 2) (i.e., approximately 3 days with 2 overnight stays). At the site's discretion, subjects could remain in the CRU for the duration of the Treatment Phase (until approximately 24 hours after the last IMP administration). Following an overnight fast, subjects received single oral doses of each of the following 7 treatments in a randomized, double-blind, crossover manner:

- GWP42003-P 750 mg
- GWP42003-P 1500 mg
- GWP42003-P 4500 mg
- Alprazolam 2 mg
- Dronabinol 10 mg
- Dronabinol 30 mg
- Placebo

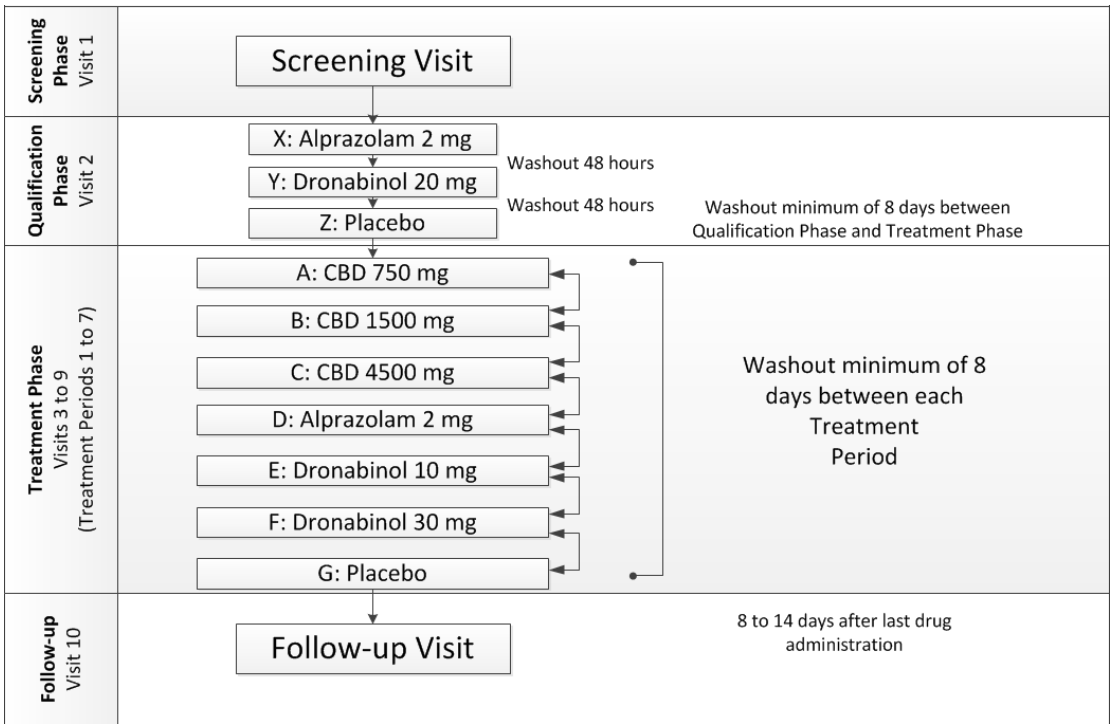
Each drug administration was separated by at least 8 days. Serial pharmacodynamic evaluations were conducted up to 24 hours after each IMP administration. Pharmacokinetic samples were obtained to confirm exposure to CBD and assess pharmacokinetic parameters. Safety monitoring included recording of AEs and regular assessments of vital signs, clinical laboratory assessments, 12-lead electrocardiogram (ECG), Columbia-Suicide Severity Rating Scale (C-SSRS) and continuous pulse oximetry/telemetry monitoring for at least 12 hours after each IMP administration.

Subjects returned for the safety Follow-up visit approximately 8 to 14 days following last drug administration.

Each subject participated in the trial for approximately 15 weeks, from Screening to Follow-up.

An overview of the Trial Schema is provided in the following figure:

Trial Schema



Note: the sequence of treatments shown is for illustration of the overall design and does not represent an actual treatment sequence.
CBD = cannabidiol (GWP42003-P).

Pharmacodynamic Endpoints:

The Emax on the bipolar Drug Liking VAS was the primary pharmacodynamic endpoint.

The secondary pharmacodynamic endpoints were as follows:

• Balance of effects:

o Drug Liking VAS (maximum effect at any dose [EmaxD], minimum effect [Emin], and time-averaged area under the effect curve to 12 hours after IMP administration [TA_AUE])

o Overall Drug Liking VAS (Emax and Emin)

o Take Drug Again VAS (Emax)

• Positive effects:

o Good Effects VAS (Emax and TA_AUE)

o High VAS (Emax and TA_AUE)

o Stoned VAS (Emax and TA_AUE)

• Negative effects:

o Bad Effects VAS (Emax and TA_AUE)

• Sedative/Stimulant effects:

o Alertness/Drowsiness VAS (Emax, Emin, and TA_AUE)

o Agitation/Relaxation VAS (Emax, Emin, and TA_AUE)

• Other drug effects:

o Any Effects VAS (Emax and TA_AUE)

o Hallucinations VAS (Emax and TA_AUE)

o Bowdle VAS (Emax and TA_AUE for internal and external perceptions sub-scales)

o Drug Similarity VAS (score at 12 hours postdose)

• Cognitive and psychomotor effects:

o Digit Symbol Substitution Test (DSST), Hopkins Verbal Learning Test - Revised (HVLT-R), and Divided Attention Test (DAT) (change from baseline to maximum/minimum effect CFBmax and/or CFBmin, and TA_AUE)

All VAS were scored on a 100-point scale. The VAS may have been administered as bipolar or unipolar scales, as appropriate, and the choice was determined by the nature of the subjective effect being measured. When VAS were administered as bipolar scales, the neutral point equaled 50 (e.g., Drug Liking, Overall Drug Liking, Alertness/Drowsiness, Agitation/Relaxation VAS). The neutral point was also labeled with an anchor, such as “neither like nor dislike.” When VAS were administered as unipolar scales, the neutral point equaled 0, and anchors were presented using text such as “Not at all” (score = 0) to “Extremely” (score = 100; e.g., Stoned, Hallucinations, Good, Bad and Any Effects VASs).

2.1.3 Number of subjects (Planned and Analyzed):

Forty-two subjects were planned to be randomized into the Treatment Phase, including 14 female subjects (one complete randomization block), to ensure that a minimum of 35 subjects completed the planned treatments (at least 1 completer per sequence).

A total of 43 subjects were randomized into the Treatment Phase, including 12 female subjects. A total of 35 subjects completed the planned treatments and were included in the pharmacodynamic analysis.

2.1.4 Pharmacodynamic Statistical Methodology used in Sponsor's analyses

Pharmacodynamic data were analyzed for the Completer Population as the primary analysis. A supplemental analysis using the primary endpoint (Drug Liking VAS Emax) may have been performed using the Per Protocol population, if substantially different (based on a review prior to database lock and unblinding) from the Completer Population. However, as these 2 populations comprised the same subjects, this analysis was not performed.

Pharmacodynamic data (Emax and/or Emin, as appropriate) from the Qualification Phase were summarized for the Completer Population (i.e., the pharmacodynamic analysis population for the main trial), using standard descriptive statistics by treatment. The data were evaluated to confirm that an appropriate population was selected for the Treatment Phase.

During the Treatment Phase, pharmacodynamic values at each time point were summarized by treatment using descriptive statistics and presented graphically (as appropriate). Derived endpoints were summarized using descriptive statistics. A mixed-effects model for a crossover trial was used to compare the primary and secondary pharmacodynamic endpoints (e.g., Emax, EmaxD, Emin, CFBmax, CFBmin, TA_AUE, where applicable) between treatments. The model included treatment, period, sequence and first-order carryover effect as fixed effects, and subject nested within treatment sequence as random effect. Baseline (predose) (where applicable) and sex were included as covariates. If the carryover effect was found to be non-significant at the 25% level, then the term was dropped from the model.

For each parameter, the above mixed-effects model was first employed, and the residuals from the model were investigated for normality using the Shapiro-Wilk W-test. If the mixed-effects model was converged and the probability value was $\geqslant$ 0.05 for the normality test, then the significance of the carryover term was checked at the 25% level. If the carryover effect was found to be non-significant at the 25% level, then the term was dropped from the analysis model. Only if the overall treatment effect was significant in the mixed-effects model, the results of stepwise pairwise comparison were displayed. If the mixed-effects model wasn't converged or the probability value was < 0.05 for the normality test, the parameter was analyzed non-parametrically. For non-parametric analysis, overall treatment effect was assessed using Friedman's test. If the overall treatment effect was significant in the Friedman's test, pairwise treatment comparisons were assessed using the following tests.

• If the difference between 2 treatments was normal distribution (i.e., the p-value was $\geqslant$ 0.05 for Shapiro-Wilk test), the paired t-test was used. Mean and 95% CIs of the differences and the p-value were presented for the pairwise differences.

• If the difference between 2 treatments was not normal distribution (i.e., the p-value was < 0.05 for Shapiro-Wilk test) but not too skewed (i.e., the skewness value was in the range of [-1, 1], the paired Z-test was used. Mean and 95% CIs of the differences and the p-value were presented for the pairwise differences.

• If the difference between 2 treatments was not normal distribution and the skewness value was less than -1 or greater than 1, the sign test was used. Median and the interquartile range (IQR: Q1 and Q3) of the differences and the p-value were presented for the pairwise differences.

The treatment comparisons to assess the abuse potential of GWP42003-P included the following:

• Alprazolam vs. placebo (trial validity)

• Each dose of dronabinol vs. placebo (trial validity)

• Each dose of GWP42003-P vs. placebo (absolute abuse potential)

• Each dose of GWP42003-P vs. alprazolam (relative abuse potential)

• Each dose of GWP42003-P vs. each dose of dronabinol (relative abuse potential)

If alprazolam and either dose of dronabinol were statistically different from placebo on the primary endpoint (Drug Liking VAS Emax), the trial was considered valid. If one of the doses of dronabinol did not statistically differentiate from placebo, this dose was not used in any further comparisons with GWP42003-P.

In addition to the above analysis, Drug Liking VAS EmaxD (maximum effect at any dose)37,38 was calculated for GWP42003-P and dronabinol. Additional comparisons were made using Drug Liking VAS EmaxD and other secondary endpoints for dronabinol vs. placebo, GWP42003-P vs. placebo, GWP42003-P vs. alprazolam, and GWP42003-P vs. dronabinol.

For each of the contrasts or pairwise comparisons, the null hypothesis was: there is no treatment effect difference between the tested pair, and the alternative hypothesis was: there is a treatment effect difference between the tested pair.

All statistical tests were performed using 2-tailed significance criteria using a 5% Type I error rate.

2.1.5 Sponsor's Pharmacodynamic Conclusions

• All 3 positive control treatments demonstrated significantly greater effects compared with placebo on the primary endpoint of Drug Liking VAS Emax, demonstrating the validity of the trial. Alprazolam and dronabinol (particularly at the 30 mg dose) also showed significantly greater effects than placebo on the majority of balance, positive, sedative and any effects measures compared with placebo and small but statistically significant negative and perceptual effects. On the Drug Similarity VAS, alprazolam was rated as being highly similar to depressants/benzodiazepines, while dronabinol showed a dose-dependent increase in ratings on the THC VAS. Alprazolam was associated with significant impairment on cognitive/psychomotor endpoints, while effects of dronabinol were more modest. Overall, appropriate responses were observed with the positive controls.

• GWP42003-P 750 mg was not significantly different from placebo on the primary endpoint or the majority of secondary endpoints. The higher doses levels of GWP42003-P (1500 and 4500 mg) demonstrated significantly greater effects compared to placebo on the primary endpoint, and there were some sporadic statistically significant differences on the secondary endpoints.

However, the magnitude of difference between GWP42003-P doses and placebo was markedly smaller than that observed with the positive controls and may not indicate a clinically important effect. In addition, GWP42003-P was not associated with significant perceptual effects or cognitive/psychomotor impairment effects, and was not rated as being similar to any drugs of abuse on the Drug Similarity VAS.

• All 3 GWP42003-P doses were associated with significantly lower effects than the positive controls on the primary endpoint and the majority of secondary endpoints of balance, positive, negative, sedative, any, perceptual and cognitive/psychomotor effects. Although there were a few endpoints that were not statistically different from the positive controls, this was generally with the highest dose of GWP42003-P relative to the dronabinol 10 mg dose, where effects were statistically significant versus placebo but relatively small in magnitude.

2.2 Data Location

The analysis datasets are located at

\\CDSESUB1\evsprod\NDA210365\0004\m5\datasets\gwep1431\analysis\adam\datasets

2.3 Reviewer's Assessment

All analyses were conducted from the stand point of the pharmacodynamics analysis.

2.3.1 Descriptive Statistics

The descriptive statistics of E_{max} and TE_{max} for the primary PD endpoint Drug Liking, and secondary PD endpoints, High, Overall Drug Liking and Take Drug Again are provided in Table 1 and Table 2. E_{max} is calculated as the maximum effect in the first 24 hours in the review's analysis. Table 1 summarizes the mean, standard deviation, minimum, the first quartile (Q1), median, the third quartile (Q3), and maximum of E_{max} for the seven treatments in the study.

Table 1. E_{max} Descriptive Statistics for Drug Liking, High, Overall Drug Liking and Take Drug Again, PD population (N=35)

Parameter	Treatment	Mean	Std Dev	Min	Q1	Median	Q3	Max
Drug Liking	A: CBD 750 mg	56.83	13.81	50.00	50.00	51.00	56.00	100.00
	B: CBD 1500 mg	61.11	16.52	50.00	50.00	51.00	69.00	100.00
	C: CBD 4500 mg	64.11	17.14	50.00	50.00	56.00	75.00	100.00
	D: Alprazolam 2 mg	79.11	15.66	51.00	66.00	79.00	94.00	100.00
	E: Dronabinol 10 mg	73.46	19.18	50.00	51.00	72.00	98.00	100.00
	F: Dronabinol 30 mg	86.74	14.82	51.00	73.00	90.00	100.00	100.00
	G: Placebo	54.63	11.14	50.00	50.00	50.00	51.00	100.00
High	A: CBD 750 mg	10.40	24.86	0.00	0.00	0.00	3.00	100.00
	B: CBD 1500 mg	20.40	34.61	0.00	0.00	0.00	25.00	100.00
	C: CBD 4500 mg	30.51	37.97	0.00	0.00	0.00	63.00	100.00
	D: Alprazolam 2 mg	55.43	38.18	0.00	21.00	65.00	94.00	100.00

	E: Dronabinol 10 mg	38.29	40.09	0.00	0.00	45.00	69.00	100.00
	F: Dronabinol 30 mg	72.51	32.92	0.00	56.00	85.00	100.00	100.00
	G: Placebo	8.71	22.27	0.00	0.00	0.00	1.00	87.00
Overall Drug Liking	A: CBD 750 mg	54.57	15.55	20.00	50.00	50.00	54.00	100.00
	B: CBD 1500 mg	56.54	18.80	22.00	50.00	50.00	63.00	100.00
	C: CBD 4500 mg	59.51	25.92	0.00	50.00	51.00	77.00	100.00
	D: Alprazolam 2 mg	86.60	16.28	35.00	75.00	93.00	100.00	100.00
	E: Dronabinol 10 mg	75.09	21.22	42.00	51.00	76.00	100.00	100.00
	F: Dronabinol 30 mg	86.83	18.98	17.00	78.00	98.00	100.00	100.00
	G: Placebo	50.09	16.53	0.00	50.00	50.00	50.00	100.00
Take Drug Again	A: CBD 750 mg	19.63	31.14	0.00	0.00	0.00	28.00	100.00
	B: CBD 1500 mg	27.49	37.28	0.00	0.00	0.00	59.00	100.00
	C: CBD 4500 mg	41.54	42.27	0.00	0.00	32.00	89.00	100.00
	D: Alprazolam 2 mg	84.86	23.72	9.00	76.00	100.00	100.00	100.00
	E: Dronabinol 10 mg	64.97	39.26	0.00	35.00	81.00	100.00	100.00
	F: Dronabinol 30 mg	84.71	27.14	0.00	70.00	100.00	100.00	100.00
	G: Placebo	10.83	25.09	0.00	0.00	0.00	0.00	100.00
Good Effects VAS	A: CBD 750 mg	22.43	32.92	0.00	0.00	5.00	33.00	100.00
	B: CBD 1500 mg	28.71	37.69	0.00	0.00	4.00	59.00	100.00
	C: CBD 4500 mg	37.51	37.98	0.00	0.00	32.00	72.00	100.00
	D: Alprazolam 2 mg	76.49	24.62	10.00	61.00	81.00	100.00	100.00
	E: Dronabinol 10 mg	54.66	38.75	0.00	8.00	70.00	90.00	100.00
	F: Dronabinol 30 mg	83.00	21.46	21.00	71.00	89.00	100.00	100.00
	G: Placebo	10.54	25.25	0.00	0.00	0.00	3.00	100.00

To be consistent with sponsor's protocol, the reviewer used:
GWP42003-P: CBD 750 mg, CBD 1500 mg and CBD 4500 mg; ALP = alprazolam; DRO = dronabinol;

From Table 1, while mean Drug Liking VAS Emax values for GWP42003-P were only slightly greater than those of placebo at the 2 higher dose levels, mean Emax with alprazolam 2 mg and dronabinol 30 mg were markedly higher (⩾ 15 points compared to placebo and all doses of GWP42003-P), with an intermediate value observed for dronabinol 10 mg. Median Drug Liking VAS Emax values for GWP42003-P doses were even lower, while median scores with alprazolam 2 mg and dronabinol 10 mg doses were similar to mean scores, or in the case of dronabinol 30 mg, slightly higher.

For High and Good Effects VAS, Emax scores remained relatively low with placebo (⩽ 10.5) and were higher for dronabinol 30 mg, followed by alprazolam 2 mg and then dronabinol 10 mg. In contrast, GWP42003-P scores were numerically higher than those of placebo, but lower than those of the positive controls.

For Overall Drug Liking VAS and Take Drug Again VAS, mean Emax scores for placebo were neutral (i.e., 50 for Overall Drug Liking VAS Emax and 0 for Take Drug Again VAS Emax), while scores for alprazolam 2 mg and both doses of dronabinol were higher. In contrast, mean Overall Drug Liking VAS Emax scores for all 3 GWP42003-P doses were only slightly higher than those of placebo, while median scores were neutral (i.e., 50.0). While mean Take Drug Again VAS Emax scores for all 3 GWP42003-P doses were higher than those of placebo, particularly at the 4500 mg dose, median scores were 0.0 for 750 and 1500 mg, and 32.0 for GWP42003-P 4500 mg. In contrast, median scores for alprazolam 2 mg and dronabinol were even higher than the mean scores (e.g., at the maximum of the scale of 100.0 for alprazolam 2 mg and dronabinol 30 mg).

Figure 1. Mean Drug Liking VAS Scores over time (Completer Population, N=35)

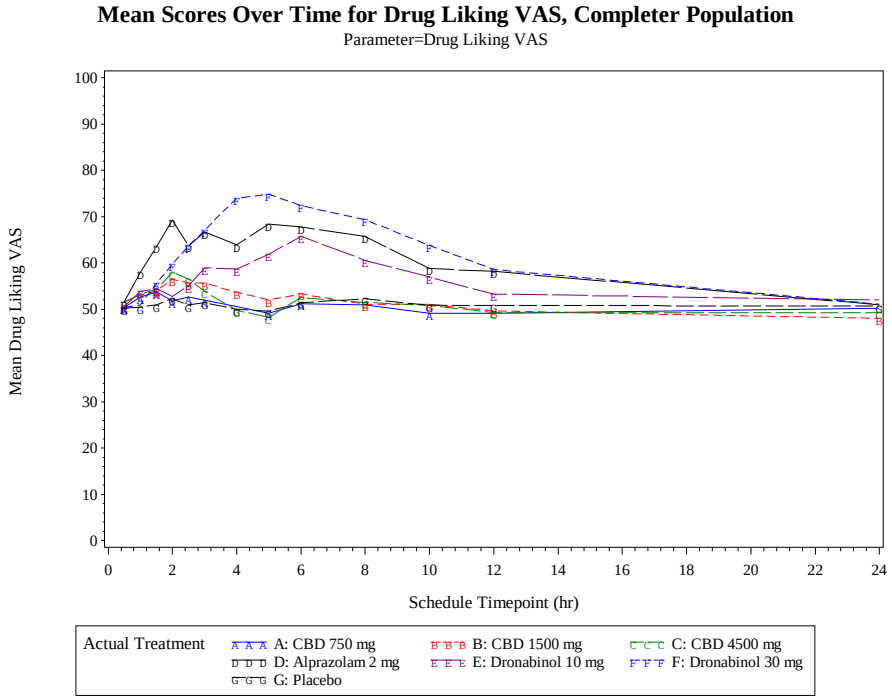


Figure 1 shows the mean drug liking VAS over time, mean scores for placebo and GWP42003-P remained within the placebo range (40 to 60, inclusive) at all timepoints, with transient marginal increases in mean scores with GWP42003-P 1500 and 4500 mg above those of placebo from approximately 1.5 to 3 or 4 hours pos-tdose (up to 52.3 at 2 hours with placebo, 53.7 at 1.5 hours with GWP42003-P 750 mg, 56.5 at 2 hours with GWP42003-P 1500 mg and 58.0 at 2 hours with GWP42003-P 4500 mg). In addition, all median scores for placebo and all 3 GWP42003-P doses were 50.0 at all timepoints, indicating that the majority of subjects neither liked nor disliked placebo and GWP42003-P. In contrast, mean Drug Liking VAS scores with alprazolam 2 mg increased up to 69.2 at 2 hours post-dose (median up to 69.0 at 5 hours), with similar but later increases with dronabinol 10 mg (mean scores up to 65.7 at 6 hours, median scores up to 59.0), while dronabinol 30 mg was associated with the largest increases in mean Drug Liking VAS scores (mean scores up to 74.9 at 5 hours, median scores up to 76.0 at 4 hours). Dronabinol and alprazolam also produced a longer duration of effects, with mean scores above the placebo range for 8 to 10 hours post-dose.

Figure 2. Mean High VAS Scores over time (Completer Population, N=35)
Mean Scores Over Time for High VAS, Completer Population
Parameter=High VAS

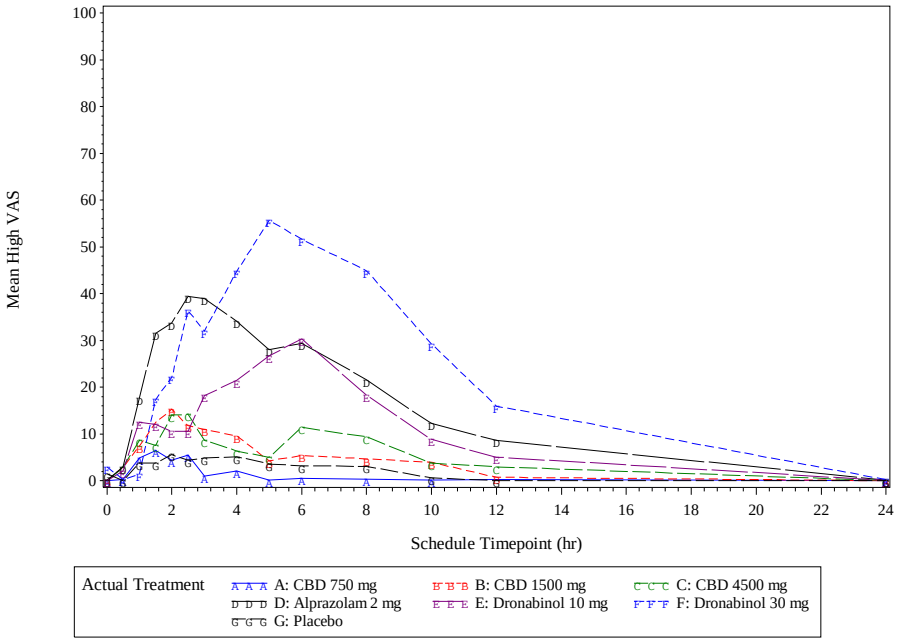


Figure 2 presented the Mean High VAS scores over time. Mean scores for placebo and GWP42003-P remained <10 at all timepoints. In contrast, mean High VAS scores with alprazolam 2 mg increased up to 40 at 2 hours post-dose, dronabinol 10 mg (mean scores up to 30 at 6 hours), while dronabinol 30 mg was associated with the largest increases in mean High VAS scores (mean scores up to 58 at 5 hours).

Figure 3. Mean High VAS Scores over time (Completer Population, N=35)
Mean Scores Over Time for Good Effect VAS, Completer Population
Parameter=Good Effects VAS

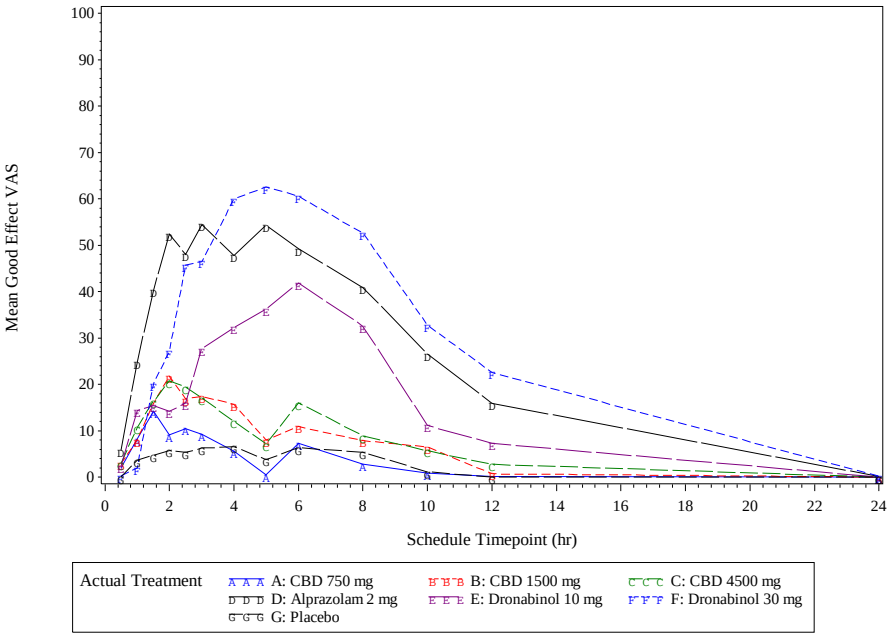


Figure 3 presented the Mean Good Effects VAS scores over time. Overall, mean Good Effects VAS scores with placebo remained relatively low (< 10), and mean scores with GWP42003-P doses were only slightly higher than those of placebo, peaking at 1.5 to 2 hours post-dose (14.2, 21.7, and 20.8 for the 3 doses, respectively) and were relatively transient, with the majority of effects lasting for only 4 to 6 hours. In addition, all median scores were 0.0 with all 3 GWP42003-P doses at all time points, indicating that the majority of subjects did not perceive any good effects with GWP42003-P. In contrast, mean scores with alprazolam 2 mg and dronabinol 30 mg increased rapidly over the first 1 to 2 hours, with mean peak effects at 3 hours with alprazolam 2 mg (up to 54.5) and 5 hours with dronabinol 30 mg (up to 62.6). Effects of both treatments lasted for at least 12 hours post-dose. Mean peak Good Effects VAS scores with dronabinol 10 mg were lower (up to 41.9) and peaked slightly later at 6 hours post-dose.

Individual E_{max} scores are displayed by subject for all treatments from Figure 4 to Figure 7, each row represents one patient with 7 treatments, the darker color means the more like. We can compare the E_{max} score for each patient at different treatment. The heatmaps show general more like for alprazolam 2 mg, dronabinol 10 mg and dronabinol 30 mg comparing with GWP42003-P, For Drug Liking VAS, some subjects had high placebo response, there were 5 out of 35 (14%) subjects had placebo response >60. In Addition, **Subject ID** (b) (6) **had placebo response=100 for all PD endpoints except for High. Subject ID=** (b) (6) **tended to have higher response for different treatment for most of the PD endpoints.**

Figure 4. Heatmap for Emax of Drug Liking VAS by treatment

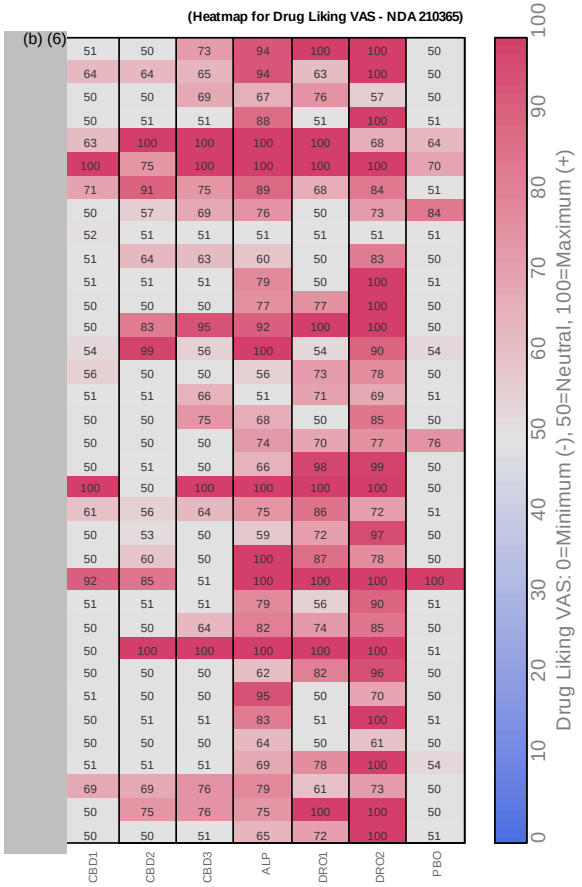


Figure 5. Heatmap for Emax of High VAS by treatment

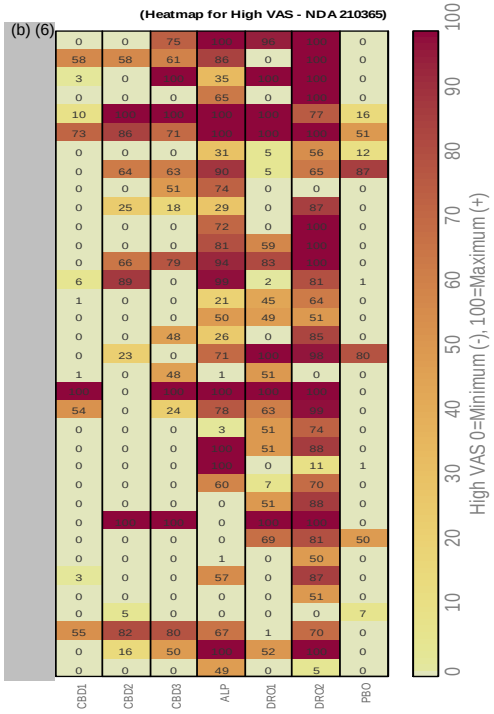


Figure 6. Heatmap for Emax of Overall Drug Liking VAS by treatment

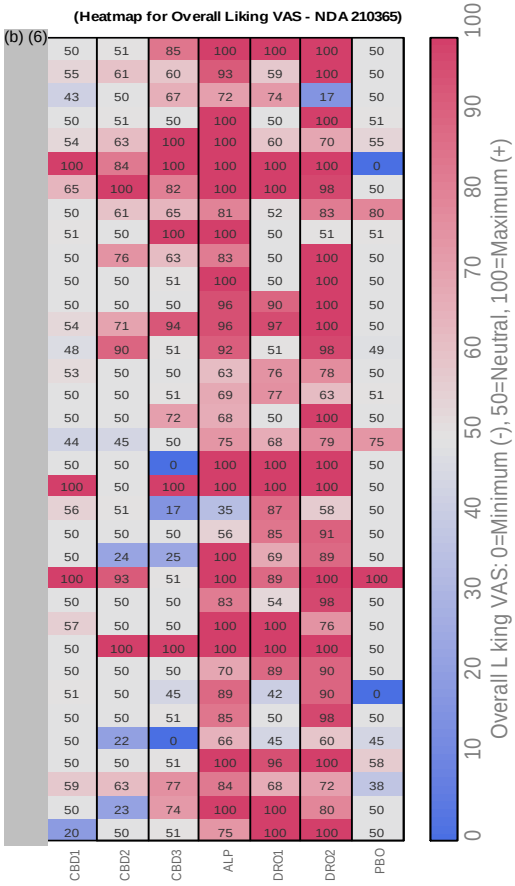


Figure 7. Heatmap for Emax of Take Drug Again VAS by treatment

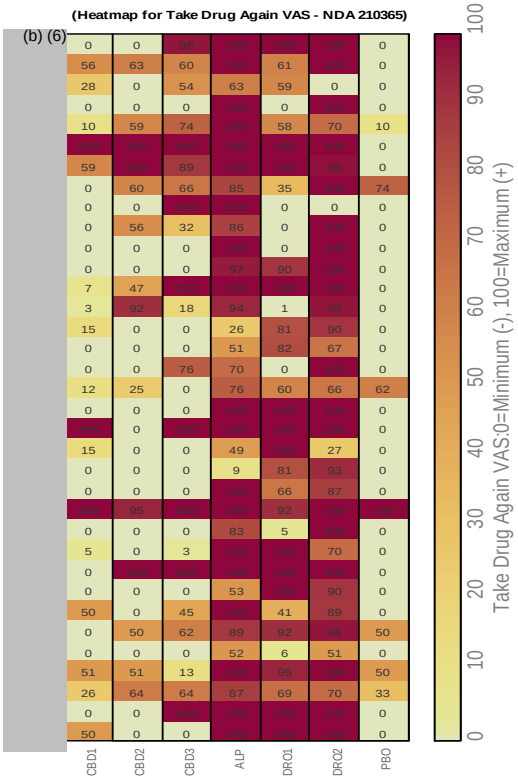
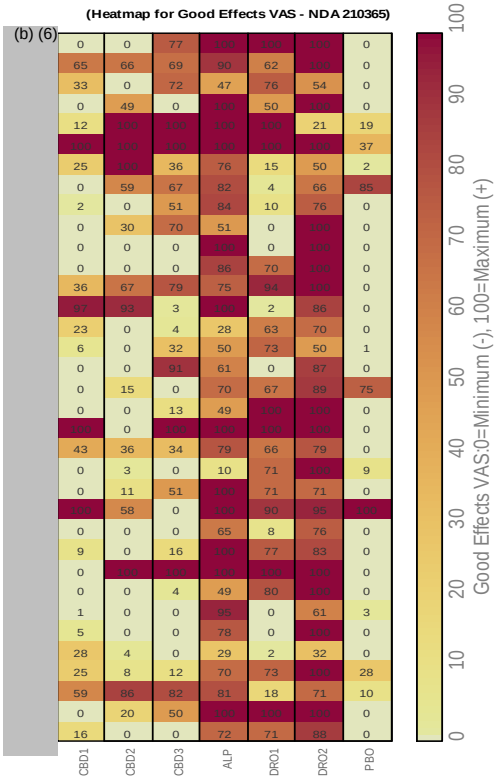


Figure 8. Heatmap for Emax of Good Effects VAS by treatment



2.3.2 Statistical Analysis

The statistical analysis of a HAP study should address whether:

• The known drug of abuse (positive control) produces reliable abuse-related responses compared to placebo.

• The test drug produces abuse- related responses that are smaller than the positive control.

• The test drug produces abuse- related responses that are similar to placebo.

To address these issues, the following hypotheses should be tested:

1. Validation of the Appropriateness of the Positive Control

(1) Primary endpoint: Drug Liking Emax

For study validity purpose, the primary endpoint, Emax for Drug Liking VAS, will be compared between each of the positive controls (Alprazolam 2 mg, Dronabinol 10 mg and Dronabinol 30 mg) and placebo. Each comparison will assess the null hypothesis that the mean difference in Drug Liking Emax between the positive control and placebo is less than or equal to 15 against the alternative hypothesis that the mean difference in Drug Liking Emax between the positive control and placebo is greater than 15. The hypothesis can be expressed as:

$$H_0: \mu_C - \mu_P \leq 15 \text{ versus } H_a: \mu_C - \mu_P > 15$$

where μ_C is the mean for the positive controls (Alprazolam 2 mg, Dronabinol 10 mg and Dronabinol 30 mg), and μ_P is the mean for placebo. If the treatment difference is statistically significant and the lower confidence limit for the difference exceeds 15, then validity is established for the study.

(2) Key Secondary endpoints: Emax for Overall Drug Liking, Take Drug Again, High, and Good Drug Effects, we used 10 as the margin. The hypothesis can be expressed as:

$$H_0: \mu_C - \mu_P \leq 10 \text{ versus } H_a: \mu_C - \mu_P > 10$$

2. Comparison between the positive controls and the test drug

Comparison between the positive controls, (Alprazolam 2 mg, Dronabinol 10 mg and Dronabinol 30 mg), and the test drug, CBD can be expressed as (where μ_T is the mean for the CBD dose):

$$H_0: \mu_C - \mu_T \leq 0 \text{ versus } H_a: \mu_C - \mu_T > 0$$

3. Comparisons between each dose of the test drug and placebo

The hypothesis for comparisons between each dose of the test drug, CBD, and placebo will be:

$$H_0: \mu_T - \mu_P \geq 11 \text{ versus } H_a: \mu_T - \mu_P < 11$$

In this study, except for Overall Drug Liking VAS, the normality assumption tests were met. The primary endpoint and key secondary endpoints will be analyzed using a mixed-effect model if the distribution of the residuals is normally distributed. The model will include treatment, period,

sequence, as fixed effects, and subject as a random effect. If this criterion is not met, each paired difference will be investigated for normality using the Shapiro-Wilk W-test. If the p-value for the distribution of the paired difference is $\geqslant$ 0.05 or the distribution is quite symmetric (skewness between -0.5 and 0.5), a paired t-test will be used. Means, SE, and one-sided 95% CIs for treatment differences will be presented. P-values will be provided for the contrasts from the paired t-tests. If the paired differences are not normally distributed and quite symmetric, pairwise treatment comparisons will be assessed using the sign test. The median, first and third quartiles, 1-sided 95% CI, and the p-value for the paired difference will be presented.

Table 2 summaries the results of Comparison of Drug Liking VAS Emax for the three tests.

Table 2. Comparison of Drug Liking VAS Emax – Primary endpoint, Completer Population

Treatments	**LS Mean**	**StdE**	**Lower**	**Upper**	
A: CBD 750 mg	58.14	2.42	53.36	62.91	
B: CBD 1500 mg	62.44	2.41	57.69	67.19	
C: CBD 4500 mg	65.55	2.42	60.78	70.33	
D: Alprazolam 2 mg	80.31	2.42	75.54	85.07	
E: Dronabinol 10 mg	74.71	2.42	69.93	79.48	
F: Dronabinol 30 mg	88.36	2.42	83.59	93.13	
G: Placebo	56.09	2.41	51.33	60.85	
Contrasts	**LS Mean**	**StdE**	**P-value**	**Lower**	**Upper**
Positive Controls vs. Placebo (Trial Validity, H_0: $\mu_C - \mu_P \leq 15$**)**					
D: Alprazolam 2 mg vs. G: Placebo	24.22	3.06	0.0015	19.15	Infty
E: Dronabinol 10 mg vs. G: Placebo	18.62	3.07	0.1198	13.55	Infty
F: Dronabinol 30 mg vs. G: Placebo	32.27	3.06	<.0001	27.22	Infty
GWP42003-P vs. Positive Controls (Relative Abuse Potential, H_0: $\mu_C - \mu_T \leq 0$**)**					
A: CBD 750 mg vs D: Alprazolam 2 mg	22.17	3.05	<.0001	17.12	Infty
B: CBD 1500 mg vs D: Alprazolam 2 mg	17.87	3.06	<.0001	12.82	Infty
C: CBD 4500 mg vs D: Alprazolam 2 mg	14.75	3.07	<.0001	9.67	Infty
A: CBD 750 mg vs E: Dronabinol 10 mg	16.57	3.06	<.0001	11.52	Infty
B: CBD 1500 mg E vs: Dronabinol 10 mg	12.26	3.06	<.0001	7.21	Infty
C: CBD 4500 mg vs E: Dronabinol 10 mg	9.15	3.08	0.0017	4.06	Infty
A: CBD 750 mg vs F: Dronabinol 30 mg	30.23	3.08	<.0001	25.13	Infty
B: CBD 1500 mg vs F: Dronabinol 30 mg	25.92	3.06	<.0001	20.87	Infty
C: CBD 4500 mg vs F: Dronabinol 30 mg	22.81	3.06	<.0001	17.76	Infty
GWP42003-P vs. Placebo (Absolute Abuse Potential, H_0: $\mu_T - \mu_P \geq 11$**)**					
A: CBD 750 mg vs G: Placebo	2.05	3.07	0.002	-Infty	7.12
B: CBD 1500 mg vs G: Placebo	6.35	3.05	0.0648	-Infty	11.40
C: CBD 4500 mg vs G: Placebo	9.46	3.05	0.3077	-Infty	14.51

Table 2 presents, for Drug Liking:

- The validity of the study was determined from the comparison of Drug Liking Emax between each positive control and placebo. The null hypothesis was defined as a mean difference in Drug Liking Emax of ⩽ 15 points between treatments. The mean difference was statistically significant for the comparisons between alprazolam and placebo, and the higher dronabinol 30 mg dose and placebo. For the lower dronabinol 10 mg dose compared with placebo, the mean difference in Emax was not statistically significant (P-value=0.1198); therefore, the criteria for study validity was not met. However, the study was designed and conducted based on the recommendations in the draft guidance on the Assessment of Abuse Potential of Drugs (Jan 2010) and completed before the final guidance was issued in Jan 2017. Thus, the study was not powered with an adequate sample size to perform this post-hoc analysis, which should be considered in the interpretation of this result.

- The relative abuse potential of GWP42003-P was evaluated by the comparison of Drug Liking Emax scores of each positive control, alprazolam and dronabinol, versus each dose of GWP42003-P. The null hypothesis was defined as a mean difference in Drug Liking Emax of ⩽ 0. All 3 GWP42003-P doses showed significantly lower Drug Liking VAS Emax scores compared with alprazolam 2 mg and both dronabinol doses ($p < 0.01$), indicating that subjects liked the positive controls significantly more than GWP42003-P.

- The absolute abuse potential of GWP42003-P was evaluated by the comparison of Drug Liking Emax between GWP42003-P and placebo. The null hypothesis was defined as a mean difference in Drug Liking Emax of ⩾ 11 points. If the null hypothesis was not rejected then the results supported that the treatments were not similar. The comparison of GWP42003-P 750 mg versus placebo was statistically significant (P value=0.002). Conversely, the comparisons of the higher GWP42003-P doses (1500 mg and 4500 mg) versus placebo were not statistically significant (P value=0.0648 and 0.3077 respectively). The results showed that responses to GWP42003-P 750 mg were similar to those for placebo, while the responses to GWP42003-P 1500 mg and 4500 mg could not be considered similar to placebo.

Since subject ID (b) (6) had placebo response=100 for all PD endpoints except for High, the reviewer also did the analysis without this subject, for the comparison GWP42003-P vs. Placebo (Absolute Abuse Potential, H0: $\mu T - \mu P \geqslant 11$), the p-value=0.0029, 0.0906 and 0.5109 respectively, even higher, conclusions are the same.

Table 3. Comparison of High VAS Emax – Completer Population

Treatments	**LS Mean**	**StdE**	**Lower**	**Upper**	
A: CBD 750 mg	14.01	5.77	2.64	25.38	
B: CBD 1500 mg	23.83	5.74	12.51	35.16	
C: CBD 4500 mg	33.57	5.77	22.20	44.94	
D: Alprazolam 2 mg	58.69	5.75	47.34	70.03	
E: Dronabinol 10 mg	40.43	5.76	29.06	51.79	
F: Dronabinol 30 mg	76.47	5.76	65.11	87.84	
G: Placebo	11.88	5.75	0.54	23.21	
Contrasts	**LS Mean**	**StdE**	**P-value**	**Lower**	**Upper**
Positive Controls vs. Placebo (Trial Validity, H_0: $\mu_C - \mu_P \leq 10$**)**					

D: Alprazolam 2 mg vs. G: Placebo	46.81	6.69	<.0001	35.76	Infty
E: Dronabinol 10 mg vs. G: Placebo	28.55	6.69	0.003	17.50	Infty
F: Dronabinol 30 mg vs. G: Placebo	64.60	6.68	<.0001	53.57	Infty
GWP42003-P vs. Positive Controls (Relative Abuse Potential, H_0: $\mu_C - \mu_T \leq 0$**)**					
A: CBD 750 mg vs D: Alprazolam 2 mg	44.67	6.67	<.0001	33.66	Infty
B: CBD 1500 mg vs D: Alprazolam 2 mg	34.86	6.67	<.0001	23.84	Infty
C: CBD 4500 mg vs D: Alprazolam 2 mg	25.12	6.71	0.0001	14.03	Infty
A: CBD 750 mg vs E: Dronabinol 10 mg	26.41	6.67	<.0001	15.39	Infty
B: CBD 1500 mg vs E: Dronabinol 10 mg	16.60	6.68	0.0069	5.56	Infty
C: CBD 4500 mg vs E: Dronabinol 10 mg	6.86	6.72	0.1544	-4.25	Infty
A: CBD 750 mg vs F: Dronabinol 30 mg	62.46	6.73	<.0001	51.34	Infty
B: CBD 1500 mg vs F: Dronabinol 30 mg	52.64	6.67	<.0001	41.62	Infty
C: CBD 4500 mg vs F: Dronabinol 30 mg	42.90	6.67	<.0001	31.88	Infty
GWP42003-P vs. Placebo (Absolute Abuse Potential, H_0: $\mu_T - \mu_P \geq 11$**)**					
A: CBD 750 mg vs G: Placebo	2.14	6.70	0.0938	-Infty	13.21
B: CBD 1500 mg vs G: Placebo	11.95	6.67	0.5568	-Infty	22.97
C: CBD 4500 mg vs G: Placebo	21.69	6.66	0.9449	-Infty	32.71

Table 3 shows that for High VAS:

- The null hypothesis was a difference in Emax of ⩽ 10 points for each treatment contrast. Across all 3 measures, statistically significant differences were observed for all treatment comparisons versus placebo (P value <0.01). These results indicated that subjects showed significantly greater positive effects for each of the positive controls compared with placebo.

- The differences in Emax for High was determined between each positive control, alprazolam and dronabinol (10 mg and 30 mg), and each dose of GWP42003-P. The null hypothesis was defined as a mean difference in Emax of ⩽ 0. Except for the comparison between C: CBD 4500 mg and E: Dronabinol 10 mg, there is no significant difference between these two treatments, all 3 GWP42003-P doses showed significantly lower High VAS Emax scores compared with alprazolam 2 mg and both dronabinol doses (p value < 0.01).

- The differences in Emax for High was determined for each GWP42003-P dose versus placebo. The null hypothesis was a difference in Emax of ⩾ 11 points for each treatment contrast. The differences between GWP42003-P and placebo were not statistically significant, indicating that significantly higher positive effects were observed for GWP42003-P (all doses) compared with placebo.

Table 4. Comparison of Good Effects VAS Emax, Completer Population

Treatments	LS Mean	StdE	Lower	Upper	
A: CBD 750 mg	23.85	5.54	12.93	34.77	
B: CBD 1500 mg	29.73	5.51	18.86	40.59	
C: CBD 4500 mg	38.65	5.54	27.73	49.56	
D: Alprazolam 2 mg	77.41	5.52	66.52	88.30	
E: Dronabinol 10 mg	55.88	5.53	44.97	66.79	
F: Dronabinol 30 mg	84.55	5.53	73.64	95.46	
G: Placebo	12.03	5.52	1.15	22.91	
Contrasts	**LS Mean**	**StdE**	**P-value**	**Lower**	**Upper**
Positive Controls vs. Placebo (Trial Validity, H_0: $\mu_C - \mu_P \leq 10$**)**					
D: Alprazolam 2 mg vs. G: Placebo	65.38	6.82	<.0001	54.12	Infty
E: Dronabinol 10 mg vs. G: Placebo	43.85	6.82	<.0001	32.58	Infty
F: Dronabinol 30 mg vs. G: Placebo	72.53	6.81	<.0001	61.28	Infty
GWP42003-P vs. Positive Controls (Relative Abuse Potential, H_0: $\mu_C - \mu_T \leq 0$**)**					
A: CBD 750 mg vs D: Alprazolam 2 mg	53.56	6.80	<.0001	42.33	Infty
B: CBD 1500 mg vs D: Alprazolam 2 mg	47.69	6.80	<.0001	36.44	Infty
C: CBD 4500 mg vs D: Alprazolam 2 mg	38.76	6.84	<.0001	27.46	Infty
A: CBD 750 mg vs E: Dronabinol 10 mg	32.03	6.80	<.0001	20.79	Infty
B: CBD 1500 mg vs E: Dronabinol 10 mg	26.15	6.81	<.0001	14.90	Infty
C: CBD 4500 mg vs E: Dronabinol 10 mg	17.23	6.85	0.0064	5.90	Infty
A: CBD 750 mg vs F: Dronabinol 30 mg	60.70	6.86	<.0001	49.37	Infty
B: CBD 1500 mg vs F: Dronabinol 30 mg	54.83	6.80	<.0001	43.58	Infty
C: CBD 4500 mg vs F: Dronabinol 30 mg	45.91	6.80	<.0001	34.66	Infty
GWP42003-P vs. Placebo (Absolute Abuse Potential, H_0: $\mu_T - \mu_P \geq 11$**)**					
A: CBD 750 mg vs G: Placebo	11.82	6.83	0.5478	-Infty	23.12
B: CBD 1500 mg vs G: Placebo	17.70	6.80	0.8372	-Infty	28.93
C: CBD 4500 mg vs G: Placebo	26.62	6.80	0.9887	-Infty	37.85

Table 4 shows that for Good Effects VAS:

- The null hypothesis was a difference in Emax of ⩽ 10 points for each treatment contrast. Across all 3 measures, statistically significant differences were observed for all treatment comparisons versus placebo (P value <0.01). These results indicated that subjects showed significantly greater positive effects for each of the positive controls compared with placebo.

- The differences in Emax for Good Effects VAS was determined between each positive control, alprazolam and dronabinol (10 mg and 30 mg), and each dose of GWP42003-P. The null hypothesis was defined as a mean difference in Emax of ⩽ 0. All 3 GWP42003-P doses

showed significantly lower Good Effects VAS Emax scores compared with alprazolam 2 mg and both dronabinol doses (p value < 0.01).

- The differences in Emax for Good Effects was determined for each GWP42003-P dose versus placebo. The null hypothesis was a difference in Emax of ⩾ 11 points for each treatment contrast. The differences between GWP42003-P and placebo were not statistically significant, indicating that significantly higher positive effects were observed for GWP42003-P (all doses) compared with placebo.

Table 5. Comparison of Take Drug Again VAS Emax, Completer Population

Treatments	**LS Mean**	**StdE**	**Lower**	**Upper**	
A: CBD 750 mg	20.61	5.96	8.87	32.36	
B: CBD 1500 mg	28.03	5.93	16.33	39.73	
C: CBD 4500 mg	41.88	5.96	30.14	53.62	
D: Alprazolam 2 mg	85.35	5.94	73.63	97.07	
E: Dronabinol 10 mg	65.59	5.95	53.85	77.32	
F: Dronabinol 30 mg	85.62	5.95	73.88	97.35	
G: Placebo	11.46	5.94	-0.25	23.17	
Contrasts	**LS Mean**	**StdE**	**P-value**	**Lower**	**Upper**
Positive Controls vs. Placebo (Trial Validity, H_0: $\mu_C - \mu_P \leq 10$**)**					
D: Alprazolam 2 mg vs. G: Placebo	73.89	6.88	<.0001	62.52	Infty
E: Dronabinol 10 mg vs. G: Placebo	54.13	6.88	<.0001	42.75	Infty
F: Dronabinol 30 mg vs. G: Placebo	74.15	6.87	<.0001	62.80	Infty
GWP42003-P vs. Positive Controls (Relative Abuse Potential, H_0: $\mu_C - \mu_T \leq 0$**)**					
A: CBD 750 mg vs D: Alprazolam 2 mg	64.74	6.86	<.0001	53.40	Infty
B: CBD 1500 mg vs D: Alprazolam 2 mg	57.32	6.86	<.0001	45.98	Infty
C: CBD 4500 mg vs D: Alprazolam 2 mg	43.47	6.90	<.0001	32.06	Infty
A: CBD 750 mg vs E: Dronabinol 10 mg	44.98	6.86	<.0001	33.63	Infty
B: CBD 1500 mg vs E: Dronabinol 10 mg	37.56	6.87	<.0001	26.20	Infty
C: CBD 4500 mg vs E: Dronabinol 10 mg	23.71	6.92	0.0004	12.28	Infty
A: CBD 750 mg vs F: Dronabinol 30 mg	65.00	6.92	<.0001	53.57	Infty
B: CBD 1500 mg vs F: Dronabinol 30 mg	57.58	6.87	<.0001	46.24	Infty
C: CBD 4500 mg vs F: Dronabinol 30 mg	43.73	6.86	<.0001	32.39	Infty
GWP42003-P vs. Placebo (Absolute Abuse Potential, H_0: $\mu_T - \mu_P \geq 11$**)**					
A: CBD 750 mg vs G: Placebo	9.15	6.90	0.3944	-Infty	20.55
B: CBD 1500 mg vs G: Placebo	16.57	6.86	0.7911	-Infty	27.91
C: CBD 4500 mg vs G: Placebo	30.42	6.86	0.9974	-Infty	41.75

Table 5 shows that for Take Drug Again VAS:

- The mean differences in Take Drug Again Emax was determined between each positive control (alprazolam and both doses of dronabinol) and placebo. The null hypothesis was a mean difference in Emax of ⩽ 10 points for each treatment contrast. All 3 positive control treatments demonstrated significantly higher in positive effects compared with placebo (P value <0.001), indicating that subjects had a significant preference to take each positive control again compared with placebo.

- The comparison was performed for Take Drug Again Emax between each positive control, alprazolam and dronabinol (10 mg and 30 mg), and each dose of GWP42003-P. The null hypothesis was defined as a mean difference in Emax of ⩽ 0. All 3 GWP42003-P doses showed significantly lower Take Drug Again Emax scores compared with alprazolam 2 mg and both dronabinol doses (p value < 0.01).

- The differences in Emax for Take Drug Again was determined for each GWP42003-P dose versus placebo. The null hypothesis was a difference in Emax of ⩾ 11 points for each treatment contrast. None of the treatment comparisons were statistically significant; therefore, the preference to take GWP42003-P again was not similar to the preference to take placebo.

Table 6. Comparison of Overall Drug Liking VAS Emax, Completer Population

Contrasts	**Mean (SE)/ Median (Q1-Q3)**	**P-value**
Positive Controls vs. Placebo (Trial Validity, H_0: $\mu_C - \mu_P \leq 10$**)**		
D: Alprazolam 2 mg vs. G: Placebo [1]	36.5 (3.98)	<0.0001
E: Dronabinol 10 mg vs. G: Placebo [1]	25.0 (4.45)	0.0009
F: Dronabinol 30 mg vs. G: Placebo [1]	36.7 (4.29)	<0.0001
GWP42003-P vs. Positive Controls (Relative Abuse Potential, H_0: $\mu_C - \mu_T \leq 0$**)**		
A: CBD 750 mg vs D: Alprazolam 2 mg [2]	35.0 (19.0 - 50.0)	<0.0001
B: CBD 1500 mg vs D: Alprazolam 2 mg [1]	30.1 (3.61)	<0.0001
C: CBD 4500 mg vs D: Alprazolam 2 mg [2]	20.0 (6.0 - 46.0)	<0.0001
A: CBD 750 mg vs E: Dronabinol 10 mg [2]	19.0 (0.0 - 40.0)	<0.0001
B: CBD 1500 mg vs E: Dronabinol 10 mg [1]	18.5 (4.34)	<0.0001
C: CBD 4500 mg vs E: Dronabinol 10 mg [1]	15.6 (5.11)	0.0022
A: CBD 750 mg vs F: Dronabinol 30 mg [2]	39.0 (13.0 - 50.0)	<0.0001
B: CBD 1500 mg vs F: Dronabinol 30 mg [2]	38.0 (9.0 - 50.0)	<0.0001
C: CBD 4500 mg vs F: Dronabinol 30 mg [2]	37.0 (6.0 - 49.0)	<0.0001
GWP42003-P vs. Placebo (Absolute Abuse Potential, H_0: $\mu_T - \mu_P \geq 11$**)**		
A: CBD 750 mg vs G: Placebo [2]	0.0 (-1.0 - 5.0)	<0.0001
B: CBD 1500 mg vs G: Placebo [2]	0.0 (-1.0 - 11.0)	0.0029
C: CBD 4500 mg vs G: Placebo [1]	9.4 (5.44)	0.3872

[1] The paired t-test was used to assess the treatment difference;
[2] The sign test was used to assess the treatment difference;

For Overall Drug liking, since the normal assumption is not met, each paired difference is investigated for normality using the Shapiro-Wilk W-test. If the p-value for the distribution of the paired difference is ⩾ 0.05 or the distribution is quite symmetric (skewness between -0.5 and 0.5), a paired t-test is used. If the paired differences are not normally distributed and quite symmetric, pairwise treatment comparisons is assessed using the sign test.

Table 6 shows that for Overall Drug Liking VAS:

- The mean differences in Overall Drug Liking Emax was determined between each positive control (alprazolam and both doses of dronabinol) and placebo. The null hypothesis was a mean difference in Emax of ⩽ 10 points for each treatment contrast. All 3 treatment comparisons were statistically significant, indicating that subjects liked each of the positive controls significantly more than placebo overall.

- The comparison was performed for Overall Drug Liking Emax between each positive control, alprazolam and dronabinol (10 mg and 30 mg), and each dose of GWP42003-P. The null hypothesis was defined as a mean difference in Emax of ⩽ 0. All 3 GWP42003-P doses showed significantly lower Overall Drug Liking Emax scores compared with alprazolam 2 mg and both dronabinol doses (p value < 0.01).

- The comparisons of GWP42003-P 750 mg versus placebo and GWP42003-P 1500 mg versus placebo were statistically significant, showing that similar overall drug liking between GWP42003-P (750 mg and 1500 mg) and placebo. However, at the GWP42003-P 4500 mg dose, the difference in Overall Drug Liking Emax was not significant, indicating that GWP42003-P 4500 mg and placebo were not similar.

3. Conclusions

The primary objective of this trial was to evaluate the abuse potential of single doses of GWP42003-P compared with alprazolam, dronabinol and placebo in healthy recreational polydrug users.

The reviewer analyzed the primary PD endpoint Drug Liking, and the secondary PD endpoints: High, Good Effects, Take Drug Again and Overall Drug Liking. The results from the statistical reviewer's analyses establish that:

- The study met the validity criteria based on the statistically significant differences in Drug Liking Emax between alprazolam 2 mg versus placebo and dronabinol 30 mg versus placebo. The lowest dose of the positive control, dronabinol 10 mg, did not meet the study validity criteria using a 15-point margin for Drug Liking VAS Emax compared with placebo. However, the study was designed and conducted based on the recommendations in the draft guidance on the Assessment of Abuse Potential of Drugs (Jan 2010) and completed before the final guidance was issued in Jan 2017. Thus, the study was not powered with an adequate sample size to perform this post-hoc analysis, which should be considered in the interpretation of this result.

- All 3 GWP42003-P doses were associated with significantly lower effects than the positive controls on the primary endpoint and secondary endpoints of High, Good Effects, Take Drug Again and Overall Drug Liking (P value <0.01), with the exception of dronabinol 10 mg versus GWP42003-P 4500 mg for High (P value=0.1544).

- Overall, GWP42003-P 750 mg showed little significant and no consistent abuse potential. However, higher doses of GWP42003-P (1500 and 4500 mg) demonstrated significantly greater effects compared with placebo on the primary endpoint and secondary endpoints of High, Good Effects and Take Drug Again. **Higher doses of GWP42003-P (1500 and 4500 mg) are associated with a signal for abuse potential.**

4. References

1) Guidance for Industry: Assessment of Abuse Potential for Drugs (January 2017) http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm198650.pdf

---

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

ANNA SUN
03/26/2018

QIANYU DANG
03/26/2018

YI TSONG
03/26/2018