PROTOCOL C1061003

A PHASE 2A, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, 3-ARM, PARALLEL-GROUP STUDY TO EVALUATE THE SAFETY, TOLERABILITY, AND PHARMACODYNAMICS OF PF-06835919 ADMINISTERED ONCE DAILY FOR 6 WEEKS IN ADULTS WITH NONALCOHOLIC FATTY LIVER DISEASE

STATISTICAL ANALYSIS PLAN (SAP)

Version: 1

Date: 07-NOV-2017

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

VERSION HISTORY

This Statistical Analysis Plan (SAP) for study C1061003 is based on the protocol dated 29JUN2017.

Table 1. Summary of Major Changes in SAP Amendments

SAP Version	Change	Rationale
1	Not Applicable	Not Applicable

1. INTRODUCTION

This SAP provides the detailed methodology for summary and statistical analyses of the data collected in study C1061003. This document may modify the plans outlined in the protocol; however, any major modifications of the primary endpoint definition or its analysis will also be reflected in a protocol amendment.

1.1. Study Objectives

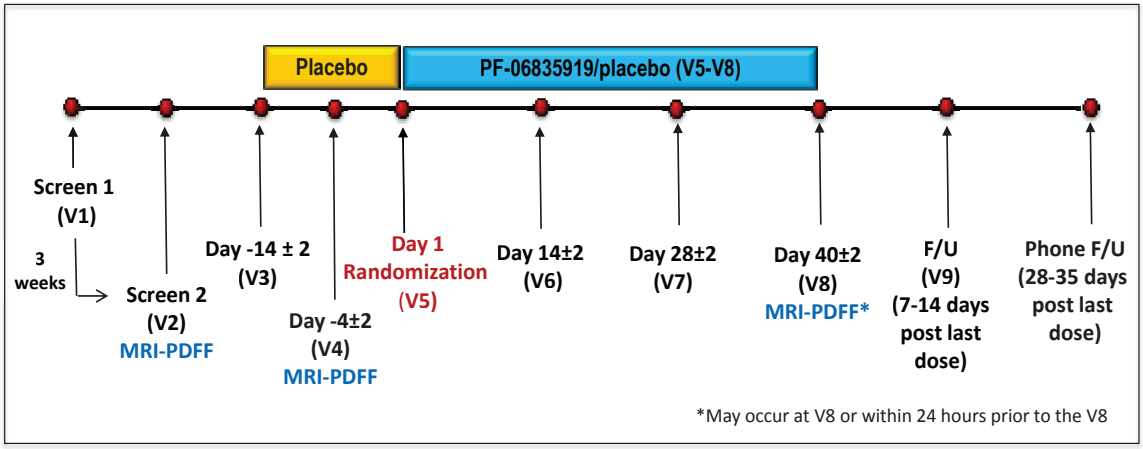
Primary Objective:	Primary Endpoint:
To evaluate the effect of daily administration of PF-06835919 over 6 weeks on liver fat in adults with NAFLD	Percent change from baseline in liver fat at Week 6, as assessed by MRI-PDFF
Secondary Objectives:	**Secondary Endpoints:**
To evaluate the safety and tolerability of PF-06835919 administered daily over 6 weeks to adults with NAFLD	Assessment of TEAEs, clinical laboratory tests, vital signs, and 12-lead ECGs
CCI	
CCI	

CCI [redacted]	[redacted]
[redacted]	[redacted]

In all cases, baseline defined as result closest ***prior to*** dosing at Visit 5 (Day 1).

1.2. Study Design

This will be a randomized, double-blind, stratified, placebo-controlled, 3-arm (placebo, plus



Determination of eligibility for this study will occur via a sequential, 2-step process, starting at the first screening visit (Screen 1). Subjects identified to be eligible based on Screen 1 procedures will proceed to Screen 2 to measure liver fat by magnetic resonance imaging-proton density fat fraction (MRI-PDFF).

Once confirmed to be eligible based on results of Screen 2, subjects will progress to a Run-in period (Visit 3), at which time subjects receive single-blind placebo for approximately 14 days to ensure compliance with the administration of IP. During the Run-in period, subjects will complete a 24-hour urine collection. In addition, approximately 4 days prior to randomization (Visit 4), subjects will report to the imaging center to have a baseline MRI-PDFF scan performed. At Visit 5 (Day 1), subjects will be randomized to receive 1 of 3 blinded IP regimens for a duration of approximately 6 weeks (ie, 40±2 days). A second 24-hour urine collection will be performed around Visit 7 (Day 28), which will require an additional visit to the study site.

This study includes a total of 9 scheduled outpatient visits to the study site, 3 visits to the imaging center, and a safety Follow-up telephone contact. The total participation, from Visit 1 (Screen 1) to the on-site Follow-up visit (Visit 9), will be up to approximately 11-14 weeks.

Up to 51 subjects (17 per arm) will be randomized at approximately 6-8 sites to ensure that a minimum of approximately 42 subjects (14 per arm) complete the study. The approximate 51 randomized subjects account for a projected premature withdrawal rate of 15%.

2. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

2.1. Primary Endpoint

Percent change from baseline in liver fat at Week 6, as assessed by MRI-PDFF. Baseline is defined as the result closest ***prior to*** dosing at Visit 5 (Day 1).

The liver fat or the whole liver PDFF is calculated from the pre-defined individual segmental PDFFs labeled Segment I, II, III, IVa, IVb, V, VI, VII and VIII as follows:

Whole Liver PDFF= PDFFs for (Segment I + Segment II + Segment III + Segment IVa + Segment IVb+Segment V + Segment VI + Segment VII + Segment VIII) / (number of segments assessed).

A minimum of 5 non-missing segments is needed in order to calculate whole liver PDFF. All segments are equally weighted.

While deriving the percent change from baseline, the same segments are to be used at both baseline and post-baseline time points in the calculation of whole liver PDFF. For example, if at baseline PDFFs from all segments are available but, at Week 6, only 7 segments have non-missing results, whole liver PDFF will be calculated using the matching individual segmental PDFFs at ***both*** baseline and Week 6.

Supplement to the primary endpoint, the percent change from baseline in each of the 9 segmental PDFFs at Week 6 will also be analysed. Baseline is defined as the result closest ***prior to*** dosing at Visit 5 (Day 1).

2.2. Secondary Endpoints

The secondary endpoints include the standard safety endpoints, namely, adverse events (AEs), clinical laboratory tests, vital signs (including blood pressure and pulse rate), and 12-lead electrocardiogram (ECG). These will be further described in Section 2.5.

CCI

[redacted]

CCI

[redacted]

[redacted]

2.4. Baseline Variables

Baseline variables are those collected on Day 1 prior to dosing or before Day 1. The demographic data of age, race, weight, and body mass index will be summarized by sex at birth and treatment in accordance with the sponsor reporting standards. Ethnicity will not be summarized. The number and proportion of subjects enrolled in each study site will also be presented.

In this study subjects will be stratified at randomization (Visit 5) based on the presence or absence of Type 2 Diabetes Mellitus (T2DM) and by the MRI-PDFF liver fat categories obtained during Visit 2 (Screen 2) listed below:

- Screening liver fat ≥6% and <10%;
- Screening liver fat ≥10%.

Both of these stratification variables will be used as covariates in the efficacy analysis. For baseline diabetic status, data will be obtained from the case report forms (CRFs). Even though baseline liver fat was used as a binary variable for stratification purposes, it will be used as a continuous covariate in the efficacy analysis in order to maximize the utility of information. Baseline fructose excretion will be considered to be included as a potential covariate. If there is substantial imbalance in the number of subjects enrolled in each site then study site will be considered to be included as a potential covariate. If included as a covariate sites enrolling less than 6 subjects will be pooled.

2.5. Safety Endpoints

The following data are considered in standard safety summaries (see protocol for collection days and list of parameters):

- AEs;
- laboratory data;
- vital signs data;
- ECG results.

2.5.1. Adverse Events

For serious adverse events (SAEs), the reporting period to Pfizer or its designated representative begins from the time that the subject provides informed consent, which is obtained prior to the subject's participation in the study, ie, prior to undergoing any study-related procedure and/or receiving investigational product, through and including 28 calendar days after the last administration of the investigational product.

Similarly,the time period for collecting AEs ("active collection period") for each subject begins from the time the subject provides informed consent. The AEs occurring following start of the double-blind randomized treatment will be counted as treatment emergent. The AEs occurring prior to the double-blind randomized treatment intake but during the single-blind placebo administration will be listed and summarized as baseline symptoms.

The 3-Tier approach will not be used to summarize the AEs due to the small size of the study.

2.5.2. Laboratory Data

Safety laboratory tests will be performed as described in the protocol.

To determine if there are any clinically significant laboratory abnormalities, the haematological, clinical chemistry (serum) and urinalysis safety tests will be assessed against the criteria specified in the sponsor reporting standards. The assessment will take into account whether each subject's baseline test result is within or outside the laboratory reference range for the particular laboratory parameter.

Baseline is defined as the result closest ***prior to*** dosing at Visit 5 (Day 1).

2.5.3. Vitals

Seated blood pressure and pulse rate measurements will be taken at time points detailed in the Schedule of Activities given in the protocol.

Baseline is defined as the result closest ***prior to*** dosing at Visit 5 (Day 1). The following vital signs endpoints will be determined for each subject:

- The maximum decrease and increase from baseline over all measurements taken post dose for systolic and diastolic blood pressures;
- The maximum increase and decrease from baseline over all measurements taken post dose for pulse rate.

The maximum increase from baseline will be calculated by first subtracting the baseline value from each post dose measurement to give the change from baseline. The maximum of these values over the respective period will then be selected, except in the case where a subject does not show an increase. In such an instance, the minimum decrease should be taken. Similarly, the maximum decrease from baseline will be determined by selecting the largest negative value of the changes from baseline. In cases where a subject does not show a decrease, the minimum increase should be taken.

2.5.4. ECGs

Single 12-lead ECGs will be obtained on all subjects at times detailed in the Schedule of Activities given in the protocol. Baseline is defined as the result closest ***prior to*** dosing at Visit 5 (Day 1).

The QT, QTcF, heart rate, QRS and PR will be recorded at each assessment time. If not supplied, QTcF will be derived using Fridericia's heart rate correction formula:

$$QTcF = QT / (RR)^{1/3} \qquad \text{where } RR = 60/HR \text{ (if RR is not provided)}$$

The maximum absolute value (post dose) and the maximum increase from baseline for QTcF, heart rate, PR and QRS, will be determined by study day for each subject.

The maximum increase from baseline will be calculated by firstly subtracting the baseline value from each post dose measurement to give the change from baseline. The maximum of these values over the respective period will then be selected, except in the case where a subject does not show an increase. In such an instance, the minimum decrease should be taken preserving the sign of change.

3. ANALYSIS SETS

Data for all subjects will be assessed to determine if subjects meet the criteria for inclusion in each analysis population prior to unblinding and releasing the database and classifications will be documented per standard operating procedures.

3.1. Full Analysis Set

The Full Analysis Set (FAS) is defined as all randomized subjects who received at least 1 dose of randomized treatment; subjects are assigned to the randomized treatment regardless of what treatment was received. This would be the primary analysis population for all efficacy analyses.

3.2. Per Protocol Analysis Set

This will not be used in the current study.

3.3. Safety Analysis Set

All subjects who receive at least 1 dose of investigational product post-randomization, ie, starting from V5, will be included in the safety analyses and listings. All safety endpoints will be analyzed by the treatment that the subjects actually receive (for the majority of the study duration) regardless of which treatment group they are randomized. A randomized but not treated subject will be excluded from the safety analyses. A treated but not randomized subject will be reported under the treatment actually received.

CCI [redacted]

4. GENERAL METHODOLOGY AND CONVENTIONS

The analysis will be performed after database release following last subject last visit.

4.1. Hypotheses and Decision Rules

The following null hypotheses will be tested for each endpoint.

1. PF-06835919 300 mg/d is equal in effect to placebo.
2. PF-06835919 75 mg/d is equal in effect to placebo.

The alternative hypotheses corresponding to the above null hypotheses will be the following 2-sided hypotheses:

1. The effect of PF-06835919 300 mg/d is different from the effect of placebo.
2. The effect of PF-06835919 75 mg/d is different from the effect of placebo.

The Type I error rate (α-level) used with the decision rule for the primary objective is 10% (2-sided). No adjustment for multiple comparisons will be made.

No interim analysis will be conducted in this study.

4.2. General Methods

Descriptive Statistics

Descriptive statistics, including the sample size, mean, standard deviation, median, minimum, and maximum values, will be provided for continuous endpoints. Some measures will be summarized using graphical representations by treatment and visit, where appropriate.

CCI

[redacted]

Analysis of Covariance (ANCOVA)

The ANCOVA model will be used with continuous endpoints for landmark (single time point) analyses. The model will include treatment group as fixed effects, baseline value and baseline diabetes status as a covariate. Baseline fructose excretion and its interaction with treatment will be included as the additional covariates for the analysis of primary endpoint only. If there is substantial imbalance in the number of subjects enrolled in each site then study site and the interaction of study site with treatment will be considered to be included as potential covariates for the analysis of primary endpoint only. If included as a covariate, sites enrolling less than 6 subjects will be pooled. The number of covariates may be reduced to improve model fit. Estimates of treatment effects will be assessed using LSMs and CIs. Estimates of the mean differences between each active dose and placebo at Week 6, and the corresponding 90% CI will be obtained from the model. Both comparisons of PF-06835919 doses with the placebo will be performed at a Type I error rate of 10% (2-sided). If there are major deviations from the statistical assumptions underlying this model then alternative transformations (eg, log) or non-parametric analyses may be presented. Justification for any alternative to the planned analysis will be given in the study report.

Non-Parametric Analysis

If the data have many outliers even after the log-transformation the following non-parametric analysis will be performed instead of the linear model. An outlier will be defined as any datapoint falling outside of 3.5 x standard deviations +/- the median. Additional evaluative statistics will be conducted to explore the nature of the outliers in order to determine the appropriateness of a parametric analysis.

For group medians 90% CIs will be presented. In addition the 90% CIs will also be presented for differences in group medians from placebo group median. The method of McGill, Tukey, and Larson will be employed to calculate the confidence interval for the difference in treatment group medians.

CCI [redacted]

[redacted]

4.3. Methods to Manage Missing Data

CCI [redacted] For the analysis of safety endpoints, the sponsor data standard rules for imputation will be applied. In case of biomarkers if the concentrations are above the limits of quantification such values will be truncated at the limit of quantification in all summary tables. If the concentrations are below the limits of quantification such values will be imputed by 0.5*LLQ where LLQ is the lower limit of quantification. However, in listings they will appear as reported.

CCI [redacted]

Note that summary statistics will not be presented at a particular time point if more than 50% of the data are missing.

5. ANALYSES AND SUMMARIES

5.1. Primary Endpoint: Whole Liver PDFF

5.1.1. Primary Analysis

- Analysis endpoint: Whole liver PDFF;
- Analysis time points: Week 6;
- Analysis population: Full Analysis set;
- Analysis methodology: %Change from baseline to Week 6 in whole liver PDFF will be analyzed using the ANCOVA. The model will include baseline whole liver PDFF, diabetes status. Baseline fructose excretion and its interaction with treatment will be considered to be included.

Reporting Results:

- Raw data: The sample size, mean, standard deviation, median, minimum and maximum at baseline and Week 6 visit will be presented for each treatment arm;
- %Change from baseline (%CBL): The sample size, mean, standard deviation, median, minimum and maximum will be presented for each treatment arm at Week 6. The LSMs, 90% CI for the LSMs, difference between the LSM of each treatment group and the placebo group, and the corresponding 90% CI and the 2-sided p-values will be presented for whole liver PDFF only.

Figures

- Whole liver PDFF - A line plot of the model-derived least square means for all treatment groups including the placebo group with 90% CI will be provided for Week 6 with dose on the X-axis;

- Whole liver PDFF - Box and whisker plots for individual percent change from baseline versus treatment will be presented and overlaid with arithmetic means.

5.1.2. Sensitivity/Robustness Analyses

- Analysis endpoint: Whole liver PDFF;

- Analysis time points: Week 6;

- Analysis population: Full Analysis set;

- Analysis methodology: Natural log-transformed individual relative change (RC) from baseline to Week 6 in whole liver PDFF will be analyzed using the ANCOVA. Log-transformed baseline will be the covariate. The model will include natural log-transformed baseline whole liver PDFF, diabetes status. Natural log-transformed baseline fructose excretion and its interaction with treatment will be considered to be included.

Reporting Results:

The LSMs and their 90% CI will be exponentiated to provide estimates of the RC which will be converted to percent change as follows:

Percent change = 100* (RC - 1).

The LSMs of the percent change from baseline with the 90%CIs for each treatment group and the LSMs of the percent change from placebo with the 90% CIs for each of the PF-06835919 groups and the 2-sided p-values will be presented.

5.1.3. Supplementary Analysis

- **Analysis endpoint:** 9 Segmental PDFFs;

- **Analysis time points:** Week 6;

- **Analysis population:** Full Analysis set;

- **Analysis methodology:** Descriptive statistics.

Reporting Results:

- Raw data: For each of the 9 segmental PDFFs the sample size, mean, standard deviation, median, minimum and maximum at baseline and Week 6 visit will be presented for each treatment arm;

- %Change from baseline (%CBL): For each of the 9 segmental PDFFs the sample size, mean, standard deviation, median, minimum and maximum will be presented for each treatment arm at Week 6.

Figures

- For each segment, box and whisker plots for individual percent change from baseline PDFF versus treatment will be presented and overlaid with arithmetic means. This will be a panel plot with 9 panels in 1 figure where the panels correspond to the segments.

5.2. Secondary Endpoints

The analyses of standard safety endpoints will be described in Section 5.6.

CCI [redacted]

CCI [redacted]

[redacted]

CCI [redacted]

CCI [redacted]

[redacted]

CCI [redacted]

CCI [redacted]

CCI [redacted]

CCI [redacted]

CCI [redacted]

CCI

[redacted]

5.4. Subset Analyses

If the analyses with baseline diabetes status indicate significant covariate effect subset analyses by baseline diabetes status may be conducted. If any site enrolls 50% or more of the total number of subjects in the study site-specific analysis may be conducted.

5.5. Baseline and Other Summaries and Analyses

A breakdown of demographic data will be provided for age, race, weight, and body mass index. Each will be summarized by sex at birth and treatment in accordance with the sponsor reporting standards. Baseline summary of parameters of metabolic syndrome, namely, triglycerides, HDL-C, waist circumference, systolic and diastolic blood pressures, AST, ALT and whole liver PDFF will be presented for each treatment group and overall. Baseline summary of stratification factors, namely, diabetes status (presence or absence), MRI-PDFF categories (<10% or ≥10%) and baseline fructose excretion will also be provided. The number and proportion of subjects enrolled in each study site will also be presented.

Subject evaluation groups will show end of study subject disposition and will show which subjects were analyzed for pharmacodynamic (FAS) and pharmacokinetics (PK), as well as for safety (adverse events and laboratory data). Frequency counts will be supplied for subject discontinuation(s) by treatment. If the proportion of discontinuations is more than 10% then table with subject disposition by visit will be produced. Data will be reported in accordance with the sponsor reporting standards.

5.6. Safety Summaries and Analyses

A set of summary tables split by treatment will be produced to evaluate any potential risk associated with the safety and toleration of administering PF-06835919.

No formal analyses are planned for safety data. The safety endpoints detailed in Section 2.5 will be listed and summarized in accordance with sponsor reporting standards, where the resulting data presentations will consist of subjects from the safety analysis sets.

Any untoward findings identified on physical examinations conducted during the active collection period will be captured as AEs, if those findings meet the definition of an AE. Data collected at screening that are used for inclusion/exclusion criteria, such as laboratory data, ECGs, and vital signs, will be considered source data, and will not be required to be reported, unless otherwise noted.

5.6.1. Adverse Events

Adverse events will be summarized by treatment and in accordance with current Pfizer data standards. The AEs will be sorted alphabetically within a system organ class. Summary tables will be provided separately for TEAEs and baseline symptoms.

The 3-Tier approach will not be used to summarize the AEs due to the small sample size.

5.6.2. Laboratory Data

All planned, quantitative, standard safety laboratory data presented in Table 6 of the protocol and non-standard safety laboratory data that are not reported independently will be listed and summarized at each planned collection time point by treatment in accordance with the sponsor reporting standards as applicable. Summary tables will present the number of observations, median, minimum and maximum. Baseline is as defined in Section 2.5.2.

5.6.3. Vital Signs

Absolute values and changes from baseline in supine systolic and diastolic blood pressure and pulse rate will be summarized by treatment and time postdose, according to sponsor reporting standards. Tables will be paged by parameter. Baseline is as defined in Section 2.5.3.

Mean changes from baseline for systolic and diastolic blood pressure and pulse rate will be plotted against time post doses. On each plot there will be 1 line for each treatment, all treatments on the same plot including the placebo. Corresponding individual plots of changes from baseline will also be produced for each treatment.

For baseline subtracted seated systolic and diastolic blood pressure and pulse rate, the differences between each dose and placebo (dose – placebo) will be summarized (N, mean, 90% CI) and plotted (mean, 90% CI) for each dose and day.

Maximum absolute values and maximum changes from baseline for vital signs, over all measurements taken post dose will also be tabulated by treatment using categories as defined in the Appendix. Numbers and percentages of subjects meeting the categorical criteria will be provided. All planned and unplanned post dose time points will be counted in these categorical summaries. All values meeting the criteria of potential clinical concern will be listed.

Maximum decrease and increase from baseline for seated systolic and diastolic blood pressures, and maximum increase from baseline for seated pulse rate will be summarized by treatment, according to sponsor reporting standards.

Changes from baseline in vital signs will also be plotted separately against drug concentrations for each time point of collection (Weeks 2, 4 and 6). These will be scatter plots for all observations where vital signs and drug concentration are recorded. Placebo data will also be included (with drug concentration set to zero). Different symbols will be used for each treatment.

5.6.4. Electrocardiogram

Absolute values and changes from baseline in QT, heart rate, QTcF, PR, RR and QRS will be summarized by treatment and day using sponsor reporting standards. Tables will be paged by parameter. Baseline is as defined in Section 2.5.4.

Mean changes from baseline in QT, heart rate and QTcF will be plotted against time postdose. On each plot there will be 1 line for each treatment, all treatments on the same plot including the placebo. Corresponding individual plots of changes from baseline will also be produced for each treatment.

In addition for baseline subtracted QT, heart rate and QTcF, the differences between each dose and placebo (dose – placebo) will be summarized (N, mean, 90% CI) and plotted (mean) for each dose and day (Weeks 2, 4, 6 and follow-up).

Changes from baseline in QTcF will also be plotted separately against drug concentrations for each time point of collection (Weeks 2, 4 and 6). These will be scatter plots for all observations where QTcF and drug concentration are recorded. Placebo data will also be included (with drug concentration set to zero). Different symbols will be used for each treatment.

ECG endpoints and changes from baseline (QTcF, PR and QRS) will also be summarized descriptively by treatment using categories as defined in the Appendix (for QTc these correspond to ICH E14). Numbers and percentages of subjects meeting the categorical criteria will be provided. All planned and unplanned postdose timepoints will be counted in these categorical summaries. All values meeting the criteria of potential clinical concern will be listed.

Maximum absolute value (post dose) and the maximum increase from baseline for QTcF, PR and QRS will be summarized by treatment according to sponsor reporting standards.

Listings of subjects with any single post dose value ≥500 msec will also be produced for QTcF.

6. INTERIM ANALYSES

No interim analysis will be conducted in this study.

7. REFERENCES

1. McGill, R., John W. Tukey and W. A. Larsen. 1978. "Variations of Box Plots." American Statistician 32:12-16

8. APPENDIX

Categorical Classes for ECG and Vital Signs

Categories for QTcF

QTcF (ms)	450≤ max. <480	480≤ max.<500	max. ≥500
QTcF (ms) increase from baseline	30≤ max. <60	max. ≥60	

Categories for PR and QRS

PR (ms)	max. ≥300	
PR (ms) increase from baseline	Baseline >200 and max. ≥25% increase	Baseline ≤200 and max. ≥50% increase
QRS (ms)	max. ≥140	
QRS (ms) increase from baseline	≥50% increase	

Categories for Vital Signs

Systolic BP (mm Hg)	min. <90	
Systolic BP (mm Hg) change from baseline	max. decrease ≥30	max. increase ≥30
Diastolic BP (mm Hg)	min. <50	
Diastolic BP (mm Hg) change from baseline	max. decrease ≥20	max. increase ≥20
Seated pulse rate (bpm)	min. <40	max. >120

- Measurements that fulfill these criteria are to be listed in the report.