Entering Outcome Measures in ClinicalTrials.gov

Version Date: 10/29/2020 | Author: Erin McDonagh, Clinical Research Support

Contents

Introduction: Outcome Measures versus Protocol Outcomes

The most common reason a ClinicalTrials.gov record is rejected during QA review is a problem with how the Outcome Measures were entered. Often this is due to:

- Misunderstanding the differences between outcome measures and protocol aims
- Inappropriately combining outcomes
- Failing to sufficiently describe the measurement to meet QC criteria
- Problems with how the outcome measure Time Frame was entered

Understanding Aims/Objectives versus Outcome Measures

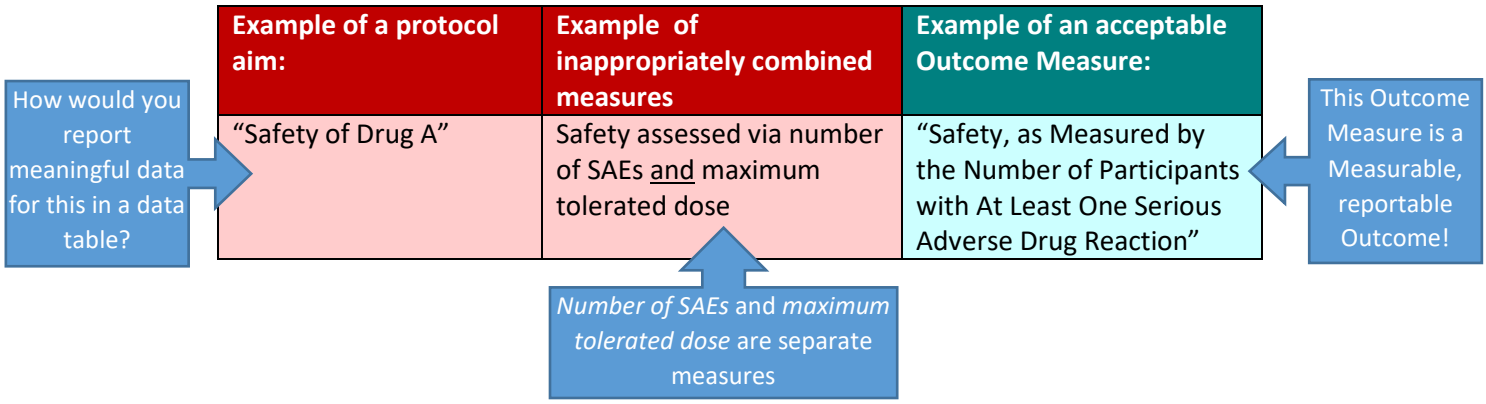
Aims and objectives are the deliverables of your study and describe what you intend to DO. They often are expressed with verbs, like "to determine…", "to assess…" etc. Outcome measures on the other hand are the individual data measurements or variables that you are collecting for your study. So, an **example of a protocol aim** might be "*To assess the safety of Drug A*", while an associated **outcome measure** might be "*Number of Serious Adverse Events*".

Aims/Objectives:	Outcome Measures
Serve as deliverables. They involve **intent to do something** with data derived from outcome measures.	**Outcome measures** are the data **measurements** gathered by the study.
Are typically expressed with **verbs.**	Outcome measures are measurements expressed in quantifiable units (with **nouns**).
Example: ***To assess*** *the efficacy of the STOMP intervention to improve opioid risk understanding and decision-making*	Example: ***Number*** *of opioid-related adverse events*

Outcome measures must be clearly described as primary, secondary, or other/exploratory in the study protocol. For most clinical trials, primary and secondary outcome measure data is required to be reported to clinicaltrials.gov 12 months after the last study visit where you collected data for that outcome measure. So, if an outcome measure is designated as primary in the protocol, it should be entered as a primary outcome in the study record. Exploratory outcome measures don't have to be reported as long as they are clearly described as exploratory in the protocol.

Protocol Outcome Measures in ClinicalTrials.gov

Protocols often summarize an amalgam of measurements under a broad outcome being evaluated by the study. Outcome Measures are *single measurable variables* used to assess that broad protocol outcome.



A good rule of thumb is that you should be able to **report a number for your outcome measure** and a public audience should be able to understand what it means. For example, try reporting a score of 10 for each of the examples above:

Unacceptable Example 1: **"Safety of Drug A" = 10**

- 10 what? 10 safeties? We don't have enough information to understand what the reported data mean.

Unacceptable Example 2: **"Safety assessed via number of SAEs and maximum tolerated dose" = 10**

- We can narrow it down to 10 meaning either 10 SAEs or 10 units of measure (like mg^m2) of the dose, but since these measurements have been inappropriately combined and have different units of measure, we still can't tell what "10" means.

Acceptable Example 3: **"Safety, as Measured by the Number of Participants with At Least One Serious Adverse Drug Reaction" = 10**

- This outcome measure clearly tells us that a score of 10 would mean 10 participants had at least one adverse drug reaction! This outcome measure is a measureable, reportable outcome and would be acceptable.

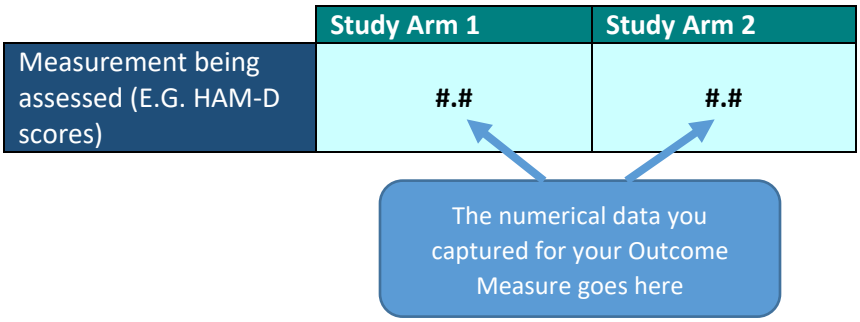
Writing Acceptable Outcome Measures:

Use this template for an Outcome Measure that includes a Score or Scale:

Template	The **[FULL NAME OF THE SCORE/SCALE]** measures **[WHAT IT MEASURES]**. Possible scores range from **[MINIMUM POSSIBLE SCORE]** to **[MAXIMUM POSSIBLE SCORE]**, with higher scores indicating a **[BETTER/WORSE]** outcome.
Example	The **Hamilton Depression Rating Scale (Ham-D)** measures **the severity of depressive symptoms**. Possible scores range from **0** to **50**, with higher scores indicating a **worse outcome**/greater severity of depressive symptoms.

For Outcome Measures that don't include a score or a scale, be sure that your Outcome Measure:

1. Is only assessing a SINGLE REPORTABLE VARIABLE
2. Is a **measurable** outcome that can be reported in a data table. Outcome Measures for most Clinical Trials must have data reported to ClinicalTrials.gov for each Primary and Secondary Outcome Measure within 12 months of the last study visit. Make sure the data for your Outcome Measure will be easy to enter into a data table that looks like this:



3. Includes enough information so that the general public can understand what is being measured and what the data mean
4. Describes *WHAT* is being measured, not *WHY* it is being measured. Avoid phrases like "To assess…" "To determine…" etc.

Outcome Measure Checklist

Refer to the checklist below to make sure your Outcome Measures will pass ClinicalTrials.gov's QA review:

Outcome Measures Check	**Complete?**
Outcome Measure Title & Description are specific	

• **Good Example:** "Pain Levels as measured by Visual Analog Score (VAS) at 6 Weeks Post-Surgery" • **Unacceptable Example:** "Discomfort After Surgery"	
Outcome Measures are listed individually - Only 1 variable is assessed per Outcome Measure • **Good Example:** "Days of Inpatient Hospitalization" • **Unacceptable Example:** "Hospital Services Usage" with a description of "days of inpatient hospitalization **and** visits to ER"	
Outcome Description describes WHAT is being measured, not WHY it is being measured • **Good Example:** "Systolic Blood Pressure at Baseline" • **Unacceptable Examples:** "To measure blood pressure"; "To determine efficacy"	
If using a score or scale, the Outcome Measure description includes: • The **full scale name** of the scale/score • **What it measures** • The **range of possible scores** • **What higher or lower scores mean** (i.e. whether a higher number = a better outcome) **Good Example:** "Outcome 1: Acute Kidney Injury as Assessed using the AKIN Scale": "The **Acute Kidney Injury Network (AKIN)** scale will be used to **assess the presence and severity of acute kidney injury (AKI)**. Stages of acute kidney injury are defined as **1, 2, or 3**, with **3 indicating the most severe AKI**."	
Outcome Time Frame specifies the **individual time point(s)** (e.g. study visits) at which the measurement is being assessed/reported, *not* the overall duration of the study. If measure is time-to-event or similar, use an estimate or maximum time that the measurement is being collected • **Good Examples:** "Baseline", "Week 2", "Up to 20 Minutes" • **Unacceptable Examples:** "Duration of study"; "Hospital admission until discharge"	
If NOT assessing a change in an Outcome Measure over time, **each time point is listed as a separate Outcome Measure** • **Good Examples:** Outcome 1: "Visual Analog Scale (VAS) Pain Score at Baseline"; Outcome 2: "Visual Analog Scale (VAS) Pain Score at Week 6" **OR** If assessing a change in the Outcome Measure over time, the title specifies "**Change in** [*Outcome Measure*]" • **Good Example:** "Change in Visual Analog Scale (VAS) Pain Score", Time Frame: "Baseline, Week 6"	
Outcome Time Frames are spelled out ("16 Weeks", not "16W")	
Outcome Measure Title, Description, and Time Frame are consistent with each other and with the rest of the record. (*e.g. if the Description says the measurement will be assessed at 3 weeks, the Time Frame specifies "3 Weeks"*)	
Unit of measure is specified (mm, kg, score, rate of…, count of participants with…, etc.)	

Outcome Measure title is outcome neutral (e.g. “change in…” vs. “decrease in…” or “improvement in…”)	
Outcome Measure titles are written in Title Case and do not end in a period	
Acronyms are expanded on the first use and symbols are spelled out (e.g. % = “percent”)	

Addressing Common Outcome Measure QC comments:

Refer to the guidance below for help addressing the most common types of Outcome Measure QA comments.

“The Measure Includes a Scale…”

If an Outcome Measure includes a scale or a score but not enough information about the scale is included, the record may be rejected with the following comment:

“**Major Issue:** The Measure includes a scale. Please provide the following scale information…”

To address this, include the following information about the scale in the Outcome Measure description:

- The **full name** of the scale (not just the abbreviation)
- **What it measures**
- **Minimum/maximum** possible scores
- **What do higher or lower scores represent**? Are higher scores better or worse, and is there a “normal” range?
- If the scale measures several things and then sums or averages them to come up with a final score, describe how scores are combined. For example, a questionnaire in which 10 tasks are rated 1-5 in difficulty, and the score is the average of all 10 responses.

Unacceptable Outcome Measure:	**Acceptable Outcome Measure**:
Title: “Depression” **Description**: “Ham-D scores”	**Title:** “Severity of Depression as Measured by the Hamilton Depression Rating Scale” **Description:** “The Hamilton Depression Rating Scale (Ham-D) is used for rating the severity of depressive symptoms. Possible scores range from 0 to 50, with higher scores indicating greater severity of depression.”

“More than one outcome measure appears to be described”

The ClinicalTrials.gov definition of “Outcome Measure” often differs from the outcomes researchers describe in their protocols. While a stated outcome in a protocol might include the assessment of several variables, an Outcome Measure described in a ClinicalTrials.gov record should be **an individual variable or measurement used to assess the protocol outcome**.

If an Outcome Measure includes multiple measurements or variables, PRS will reject the record with the following comment:

Major Issues: More than one outcome measure appears to be described.

To resolve the issue, make sure that you are reporting each distinct measurement/variable as a separate Outcome Measure. Refer to the examples below:

Unacceptable Outcome Measure:	**Acceptable Outcome Measures**:
Outcome 1: "Body composition" **Description:** "Body Mass Index (BMI) and Visceral Fat Index (as assessed by CT)."	**Outcome 1:** "Body Mass Index (BMI)" **Description:** "Body Mass Index (BMI) is a person's weight in kilograms divided by the square of height in meters. Scores between 18.5 and 24.9 indicate a healthy weight." **Outcome 2**: "Visceral Fat Index (VFA)" **Description:** "Visceral Fat Index (VFI) will be assessed using a CT scan. Scores range between 1 and 59. Scores between 1 and 12 indicate healthy levels of visceral fat. Scores of 13 and above indicate excessive, unhealthy levels of visceral fat."

Exception: if multiple measures are being combined to report a single score or number, (i.e. "Count of participants with any of X, Y or Z), then it's ok to combine those measurements:

- **Example**: "Count of participants with either a BMI of 25 or higher or a VFI of 13 or higher at baseline."

"Outcome Measure Title and Description do not appear to provide sufficient information"

If an Outcome Measure title and description do not contain enough information for the general public to understand the measurement, PRS will reject the record with the following "Major Issue" comment:

"**The Outcome Measure Title and Description do not appear to provide sufficient information to understand what will be assessed."**

To address this, **edit the Outcome Measure so that it thoroughly describes the measurement in non-technical language**.

Be sure to include:

- Information about WHAT is being measured, not WHY it is being measured. Avoid phrases like "To assess…" "To examine…" etc.
- The full name of the measurement or the measurement tool, not just the abbreviation
- The unit of measurement (i.e. mm, kg, score, rate of…, count of participants with…, etc.)
- If using a **scale or score**, include:
    - What it measures

- Minimum/maximum scores
- What do higher or lower scores represent? Are higher scores better or worse, and is there a "normal" range?
- If the scale measures several things and then sums or averages them (etc.) to come up with a final score, describe that. For example, a questionnaire in which 10 tasks are rated 1-5 in difficulty, and the score is the average of all 10 responses.

Unacceptable Outcome Measure:	**Acceptable Outcome Measure**:
Outcome 1: "Satisfaction with exercise program" **Description:** "The satisfaction of participants with their assigned exercise program"	**Outcome 1:** "Satisfaction with Exercise Program" **Description:** "Participant's satisfaction with their assigned exercise program as assessed using a 5-point Likert scale. Scores range from "very unsatisfied" (1) to "very satisfied" (5)"

Unacceptable Outcome Measure:	**Acceptable Outcome Measure**:
Outcome 1: "AKI" **Description:** "AKI will be determined via AKIN assessment"	**Outcome 1:** "Rate of Acute Kidney Injury (AKI)" **Description:** "The AKIN scale will be used to assess the presence and severity of acute kidney injury (AKI). The AKIN is a classification/staging system of acute kidney injury developed by the Acute Kidney Injury Network (AKIN) which uses changes in serum creatinine (SCr) and urine output to assess AKI. Stages of acute kidney injury are defined as 1, 2, or 3, with 3 indicating the most severe AKI."

Unacceptable Outcome Measure:	**Acceptable Outcome Measure**:
Outcome 1: "To Determine Physical Function" **Description:** "The primary outcome is to determine if the surgery improves physical function in patients."	**Outcome 1:** "Change in Physical Function as assed using the WOMAC Scale" **Description:** "Participant-reported physical function post-surgery as assessed using the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). The WOMAC assesses pain, stiffness, and physical function in patients with hip and / or knee osteoarthritis. Possible scores range from 0-96. Total score is computed by summing three subscales: pain (range 0-20), stiffness (range 0-8), and functional limitations (range 0-68), then dividing by total points possible. Higher scores indicate worse pain, stiffness, and functional limitations."

"Outcome Measure Time Frame does not appear to be specific and/or in the correct format"

ClinicalTrials.gov Outcome Measure Time Frames are the specific time points at which you are collecting data for (measuring) the outcome. If the Time Frames are not entered as a specific time points, the record will be rejected with the following comment:

"**Major Issues:** The Time Frame does not appear to be specific and/or in the correct format."

Usually, the Time Frame will simply be the **study visit time point(s)** (e.g. "Baseline" "Week 2") when you collect data for that Outcome Measure. Don't abbreviate units of time.

For Outcome Measures where traditional visit time points don't apply (e.g. time-to-event measures), enter the estimated period of time over which you are measuring the outcome in the format "**Up to [estimated maximum time]**".

Unacceptable Outcome Measure Time Frames:	Acceptable Outcome Measure Time Frames:
• "Duration of study" • "D1, D14, D30" • "Duration of participation" • "Hospitalization" • "Hospital admission until discharge"	• "Week 2" • "Baseline, 6 Weeks" • "Day 1, Day 14, Day 30" • "Up to 100 Weeks" • "Baseline, Up to 20 minutes"

If helpful, **include information on how you are defining the time frame boundaries**:

Acceptable Time Frame: "From date of randomization until the date of first documented progression or date of death from any cause, whichever came first, assessed up to 100 months"

"Outcome Measure Time Frame includes more than one time point"

A Common PRS "Major Issue" that will get a record rejected is if an Outcome Measure Time Frame:

1. Includes more than one time point AND
2. Isn't measuring a change between those time points.

In these cases, the record will be rejected with the following comment:

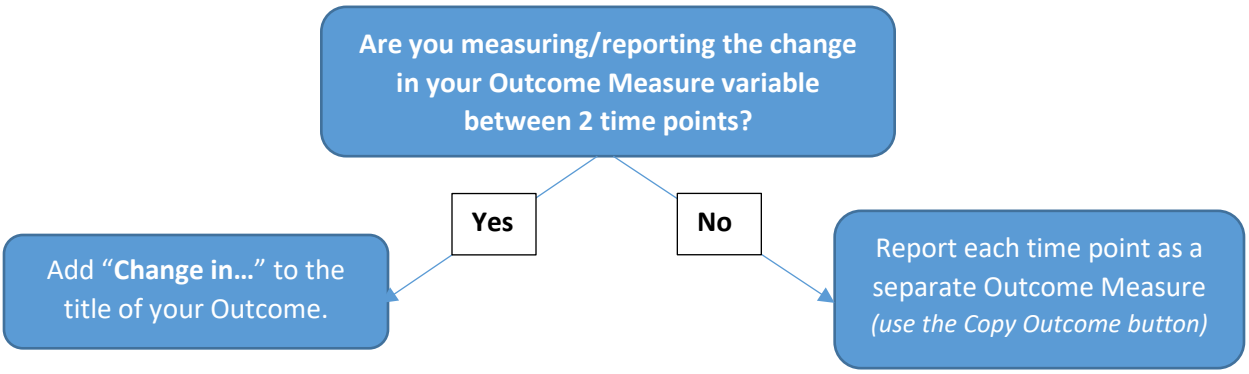
"Major Issues: The Outcome Measure Time Frame includes more than one time point, each of which appears to describe a separate measure."

- "The Outcome Measure Time Frame includes more than one time point. Each Outcome Measure should typically only specify a single time point of assessment. A common exception to this is a measure assessing change between two time points (e.g., "Change from Baseline Systolic Blood Pressure at 6 months"). If the Outcome Measure(s) are assessing a change, please revise the Outcome Measure Title(s) to specify that "change" is being assessed. If not assessing change, please revise and enter additional Outcome Measures so that there is only one Time Frame per Outcome Measure."

Resolving the issue:

First, determine whether you are measuring a CHANGE in that measurement over time, and will be reporting the change value as your Outcome Measure data.

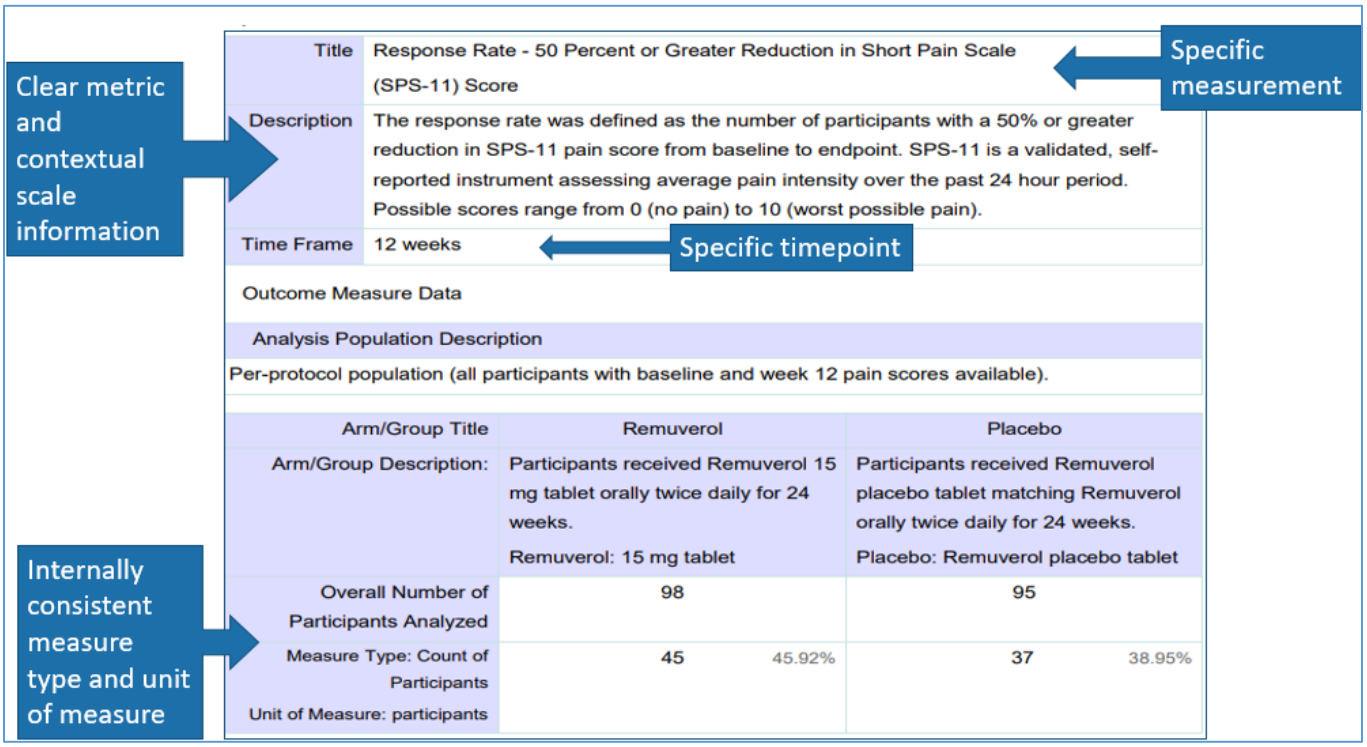
- IF you ARE measuring change in your outcome over time, add "Change in…" to the title of your outcome.
- IF you are NOT measuring change in your outcome over time, **enter each time point as a separate Outcome Measure.**
    - You can easily do this by using the "Copy Outcome" button, then adding "…at [timepoint]" to the end of the title.



Unacceptable Outcome Measure:	Acceptable, MEASURING CHANGE OVER TIME:	Acceptable, NOT MEASURING CHANGE OVER TIME:
Outcome 1: "Visual Analog Scale (VAS) Pain Score" **Time Frame:** "Baseline, 6 Weeks"	**Outcome 1:** "***Change in*** Visual Analog Scale (VAS) Pain Score" **Time Frame:** "Baseline, 6 Weeks"	**Outcome 1:** "Visual Analog Scale (VAS) Pain Score **at Baseline**" **Time Frame:** "Baseline" **Outcome 1:** "Visual Analog Scale (VAS) Pain Score **at Week 6**" **Time Frame:** "6 Weeks"

Examples of Outcome Measures with Results Data Entered

The reason the QC criteria is so strict is actually to make it easy to fit your data into the ClinicalTrials.gov outcome measure data tables when it comes time to report your results. Take a look at the three example records below to see the pre-specified outcome measure titles, descriptions, and time frames at the top, and then the generated data tables for each with outcome measure data entered.



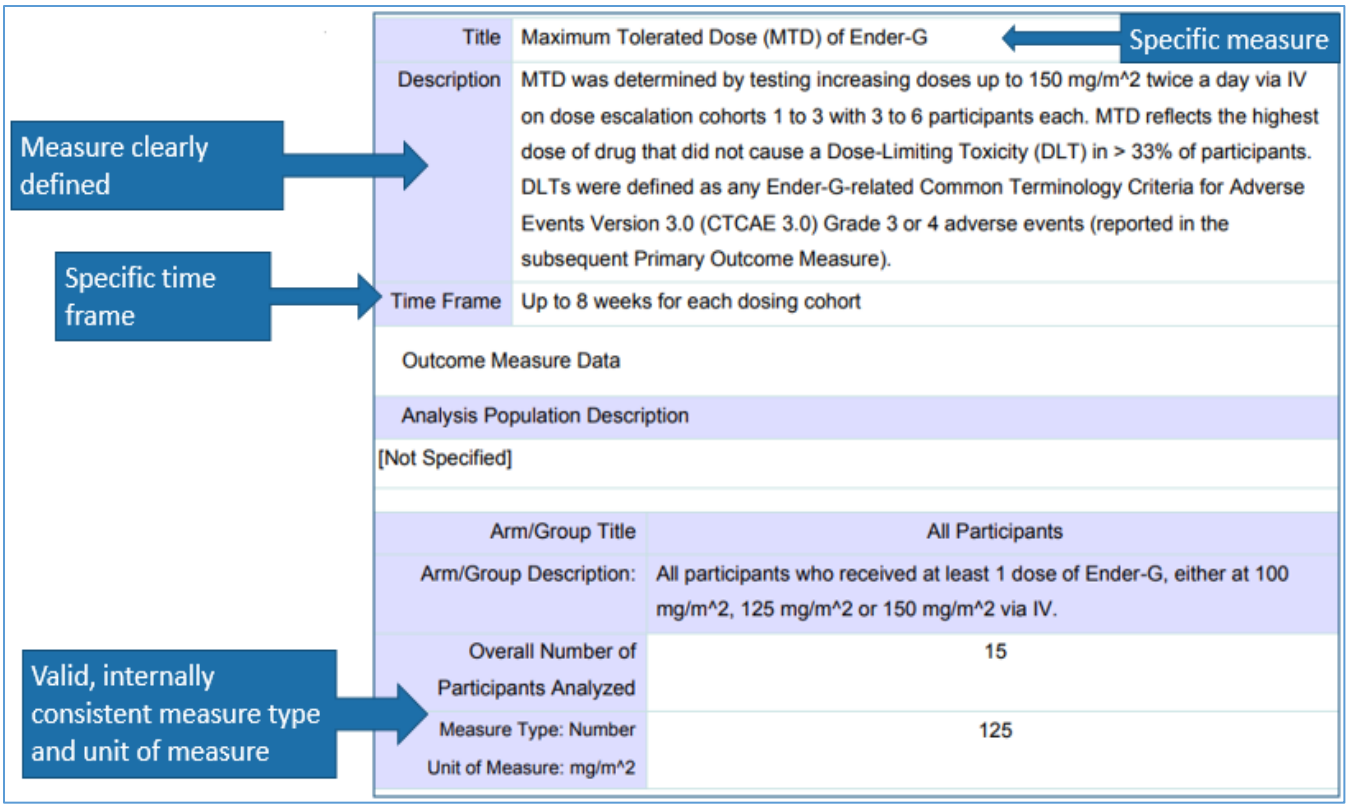
Title	Response Rate - 50 Percent or Greater Reduction in Short Pain Scale (SPS-11) Score
Description	The response rate was defined as the number of participants with a 50% or greater reduction in SPS-11 pain score from baseline to endpoint. SPS-11 is a validated, self-reported instrument assessing average pain intensity over the past 24 hour period. Possible scores range from 0 (no pain) to 10 (worst possible pain).
Time Frame	12 weeks

Outcome Measure Data

Analysis Population Description

Per-protocol population (all participants with baseline and week 12 pain scores available).

Arm/Group Title	Remuverol	Placebo
Arm/Group Description:	Participants received Remuverol 15 mg tablet orally twice daily for 24 weeks. Remuverol: 15 mg tablet	Participants received Remuverol placebo tablet matching Remuverol orally twice daily for 24 weeks. Placebo: Remuverol placebo tablet
Overall Number of Participants Analyzed	98	95
Measure Type: Count of Participants Unit of Measure: participants	45 45.92%	37 38.95%



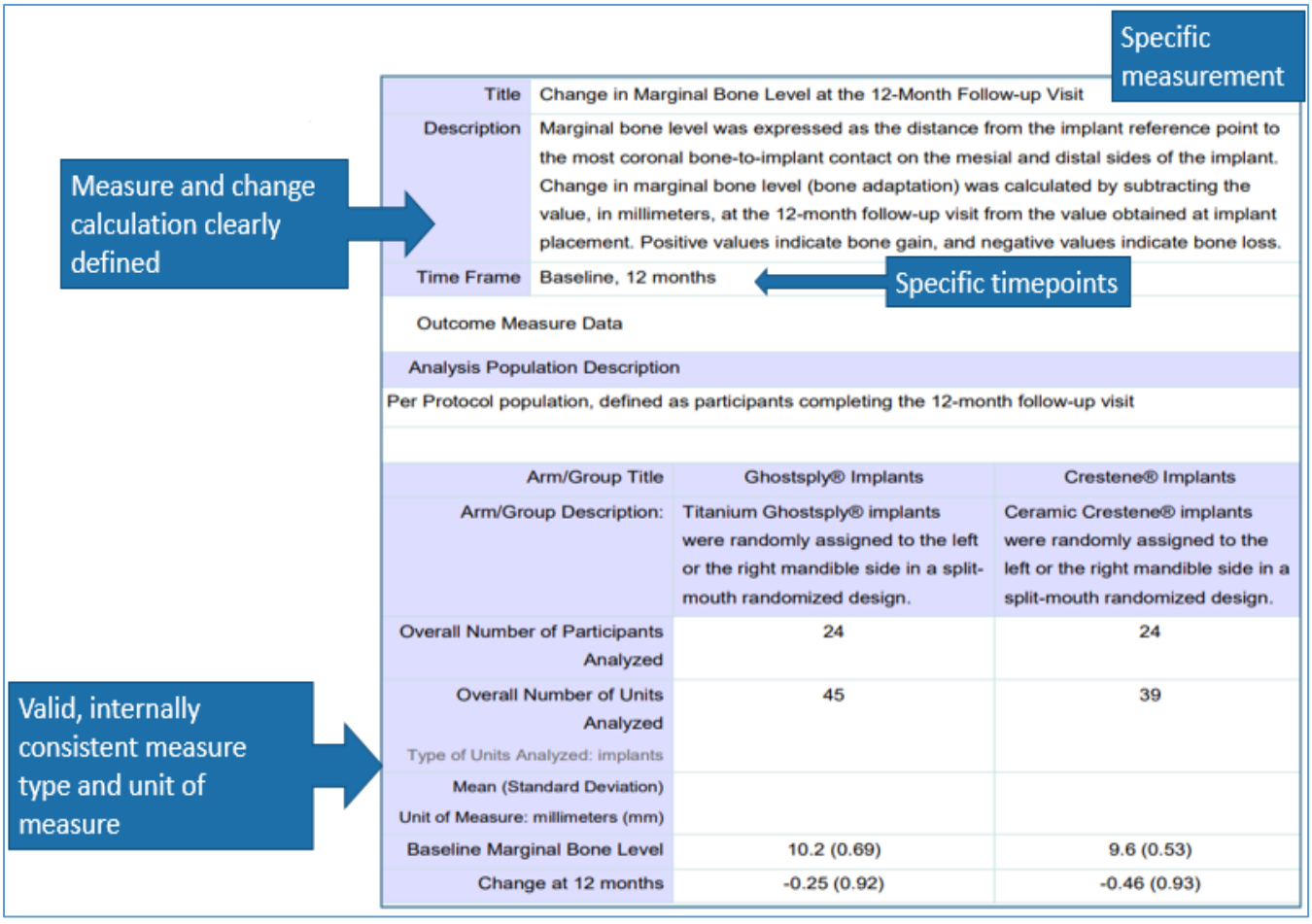
Title	Maximum Tolerated Dose (MTD) of Ender-G
Description	MTD was determined by testing increasing doses up to 150 mg/m^2 twice a day via IV on dose escalation cohorts 1 to 3 with 3 to 6 participants each. MTD reflects the highest dose of drug that did not cause a Dose-Limiting Toxicity (DLT) in > 33% of participants. DLTs were defined as any Ender-G-related Common Terminology Criteria for Adverse Events Version 3.0 (CTCAE 3.0) Grade 3 or 4 adverse events (reported in the subsequent Primary Outcome Measure).
Time Frame	Up to 8 weeks for each dosing cohort

Outcome Measure Data

Analysis Population Description

[Not Specified]

Arm/Group Title	All Participants
Arm/Group Description:	All participants who received at least 1 dose of Ender-G, either at 100 mg/m^2, 125 mg/m^2 or 150 mg/m^2 via IV.
Overall Number of Participants Analyzed	15
Measure Type: Number Unit of Measure: mg/m^2	125



Title	Change in Marginal Bone Level at the 12-Month Follow-up Visit
Description	Marginal bone level was expressed as the distance from the implant reference point to the most coronal bone-to-implant contact on the mesial and distal sides of the implant. Change in marginal bone level (bone adaptation) was calculated by subtracting the value, in millimeters, at the 12-month follow-up visit from the value obtained at implant placement. Positive values indicate bone gain, and negative values indicate bone loss.
Time Frame	Baseline, 12 months

Outcome Measure Data

Analysis Population Description

Per Protocol population, defined as participants completing the 12-month follow-up visit

Arm/Group Title	Ghostsply® Implants	Crestene® Implants
Arm/Group Description:	Titanium Ghostsply® implants were randomly assigned to the left or the right mandible side in a split-mouth randomized design.	Ceramic Crestene® implants were randomly assigned to the left or the right mandible side in a split-mouth randomized design.
Overall Number of Participants Analyzed	24	24
Overall Number of Units Analyzed Type of Units Analyzed: implants	45	39
Mean (Standard Deviation) Unit of Measure: millimeters (mm)		
Baseline Marginal Bone Level	10.2 (0.69)	9.6 (0.53)
Change at 12 months	-0.25 (0.92)	-0.46 (0.93)

Resources

Need help with ClinicalTrials.gov?

Contact Clinical Research Support at clinicalresearchsupportcenter@ucdenver.edu for help.

Visit our ClinicalTrials.gov support page, where you can find an FAQ and helpful ClinicalTrials.gov Tips of the Week.

Attend a session of CTSA31: ClinicalTrials.gov, a quarterly free lecture which covers registration, regulations, and which touches on results reporting. Upcoming sessions can be found on our Events Calendar.