

0 6 . M A R . 2 0 2 3

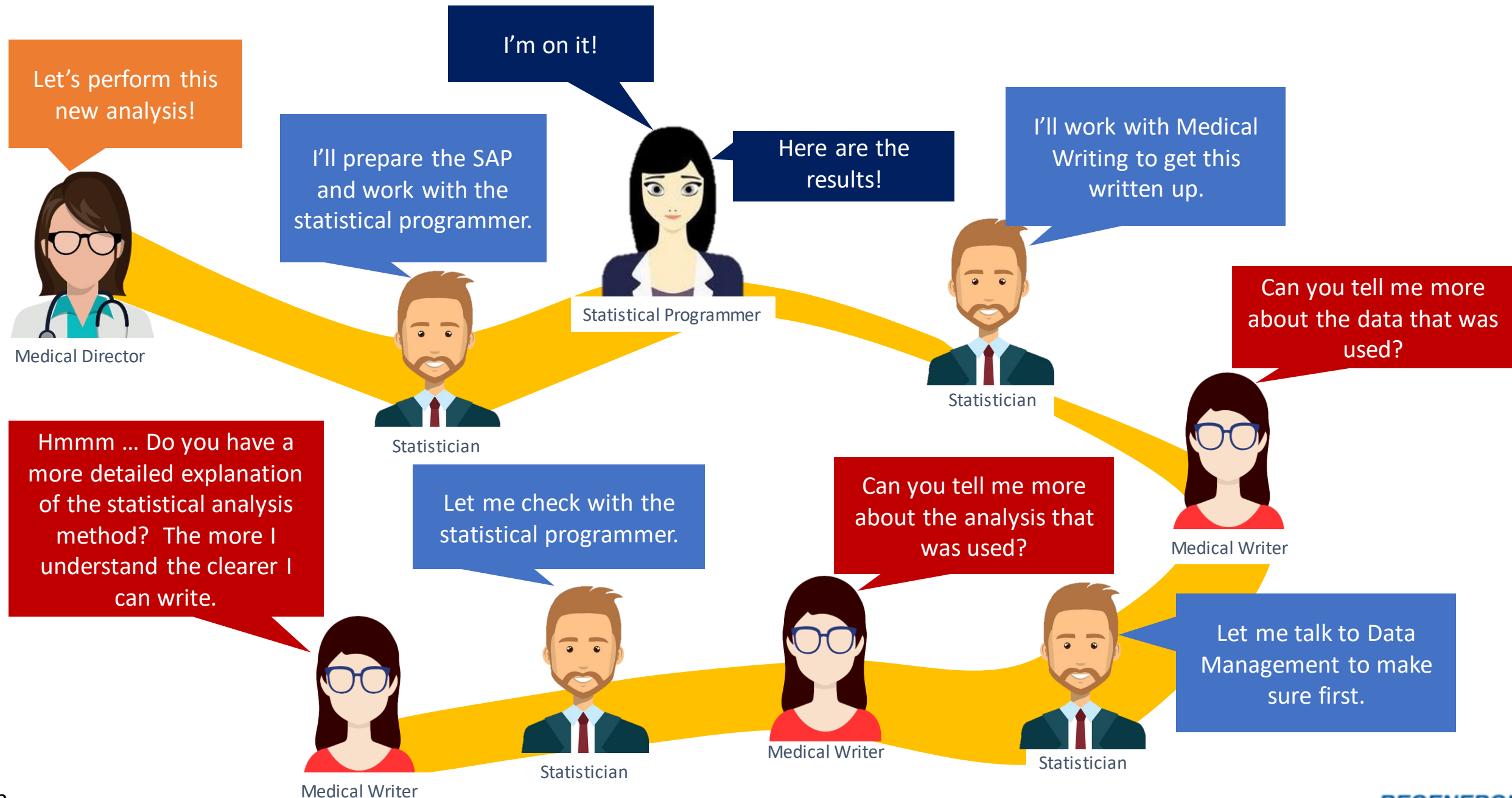
SD01: SPECification DOcument (SPECDOC)

Automating Text for Data Description, Statistical
Methods and Analysis Results for Medical
Communications

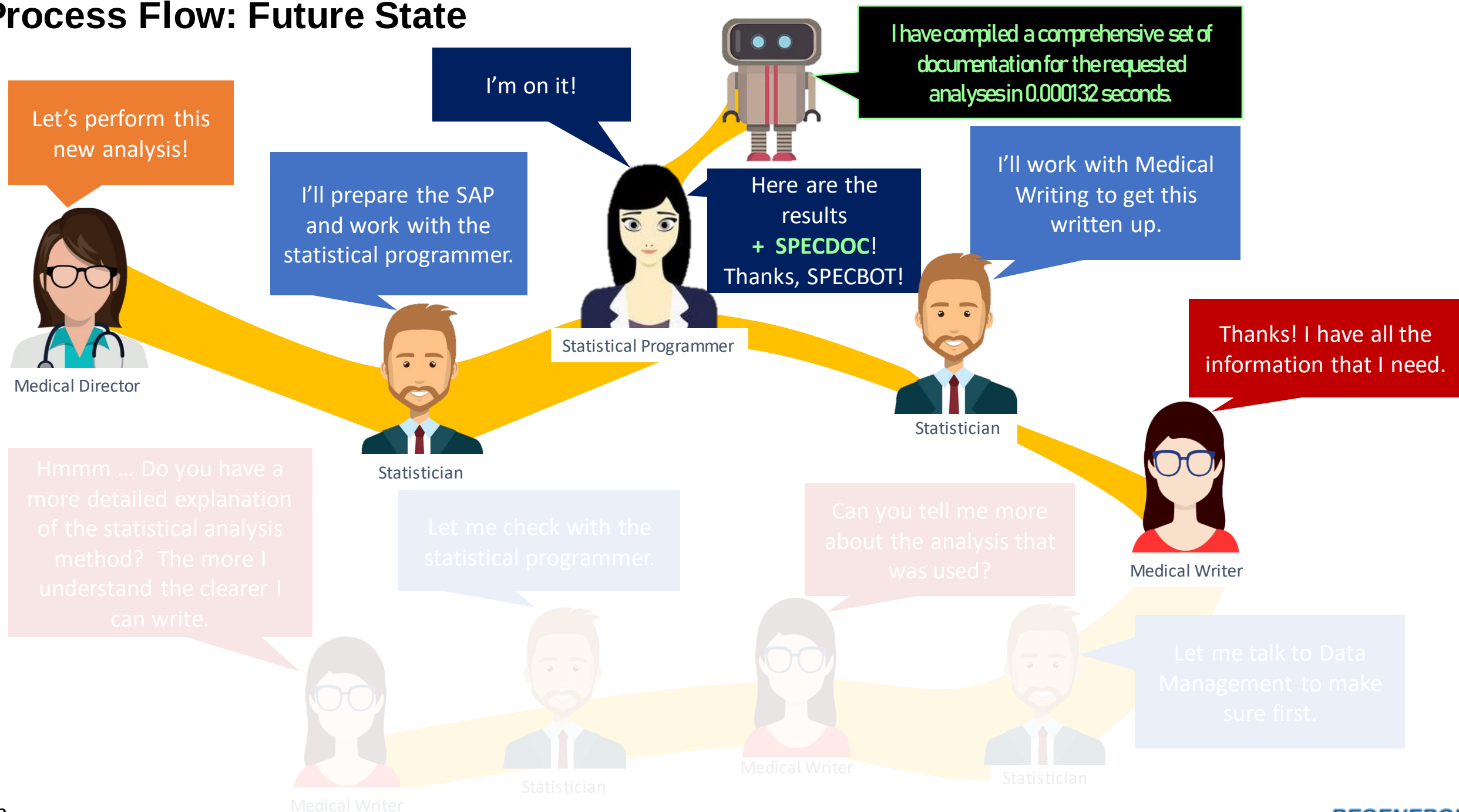
REGENERON®



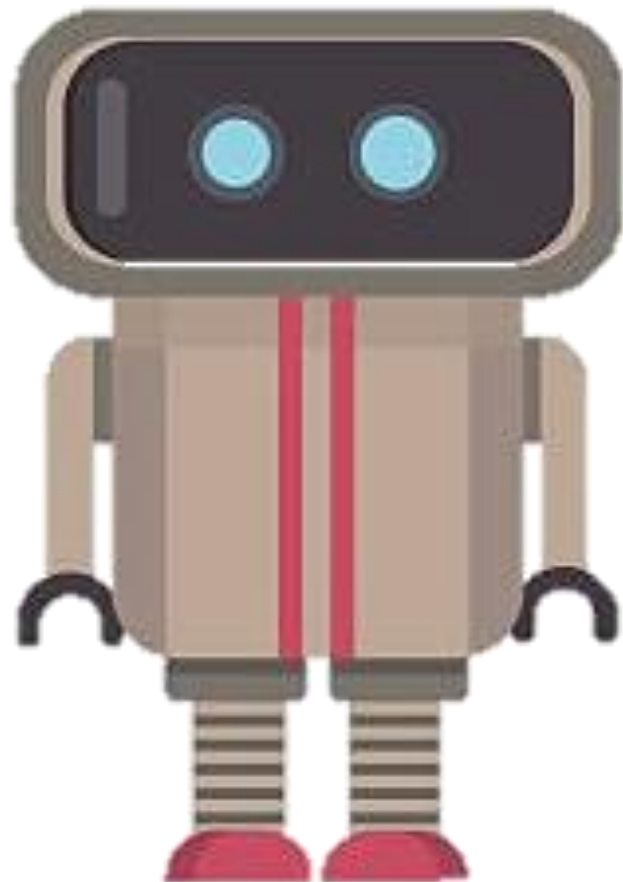
Process Flow: Current State



Process Flow: Future State



Overview



Introduction



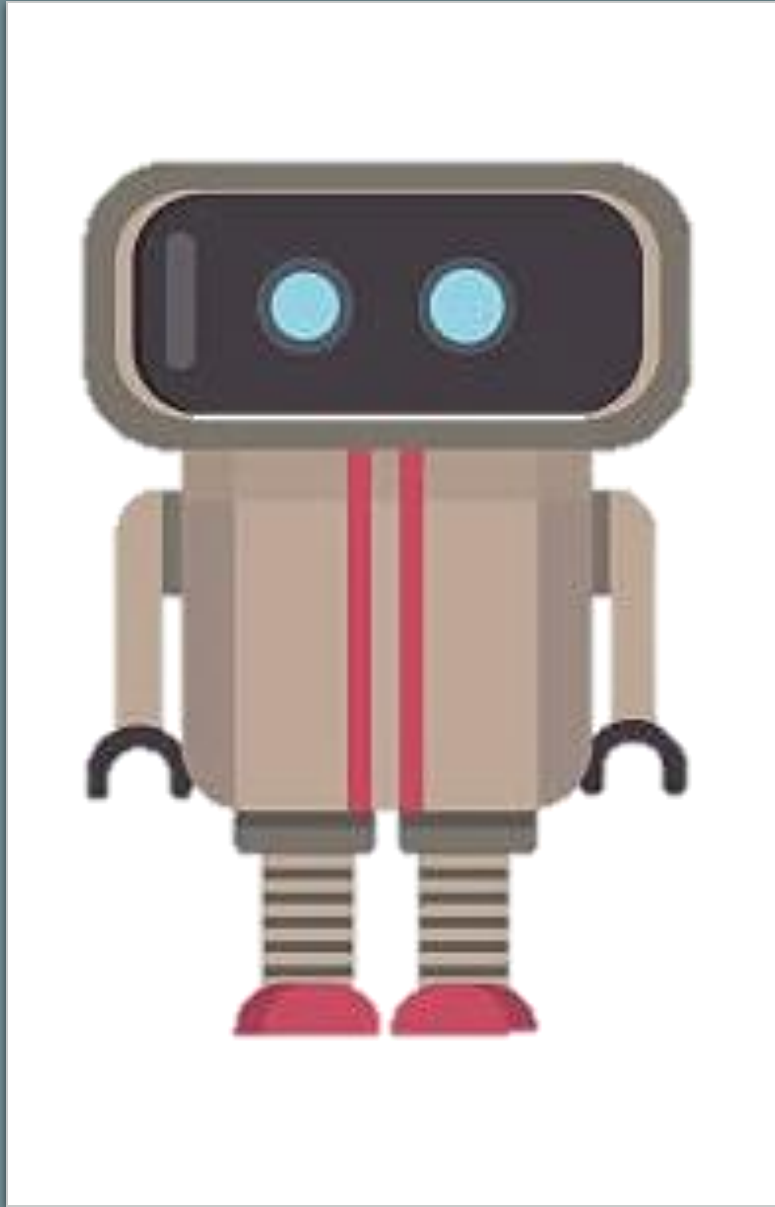
Design



Implementation



Next Steps



Overview



Introduction



Design



Implementation



Next Steps

HELLOWORLD

SPECification DOCument (SPECDOC)

Automatically generate text for data description, statistical methods, and analysis results.

I can perform the following tasks.

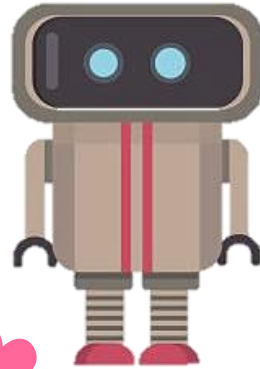
- Automatically provide detailed and comprehensive description of data and statistical analysis methods
- Automatically generate text to describe output and results
- Automatically linked to additional statistical methods resources



Introduction: Who is it for?



Medical Writer
Publications



SPECBOT



Statistician



Statistical
Programmer



Medical Directors



Medical Affairs and Beyond

Overview



Introduction



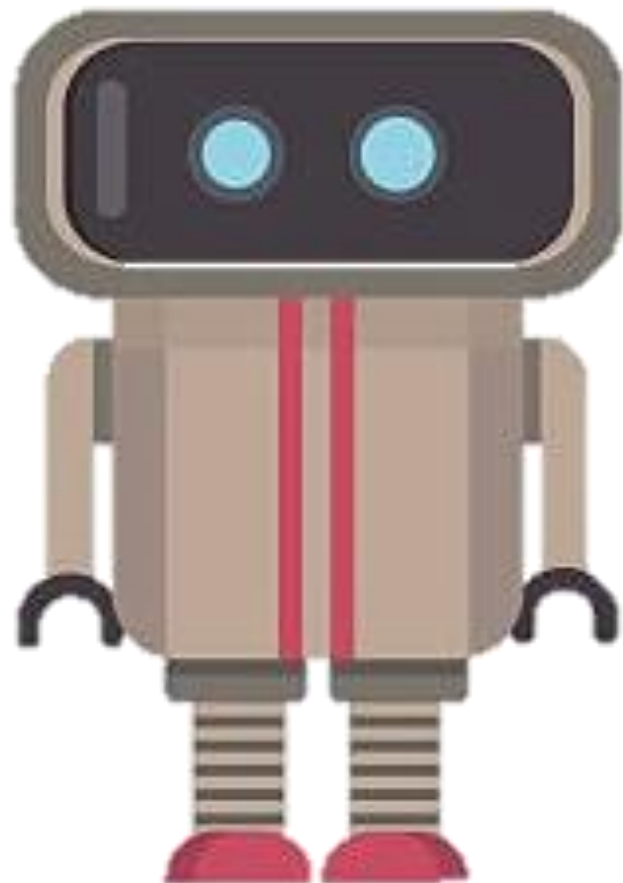
Design



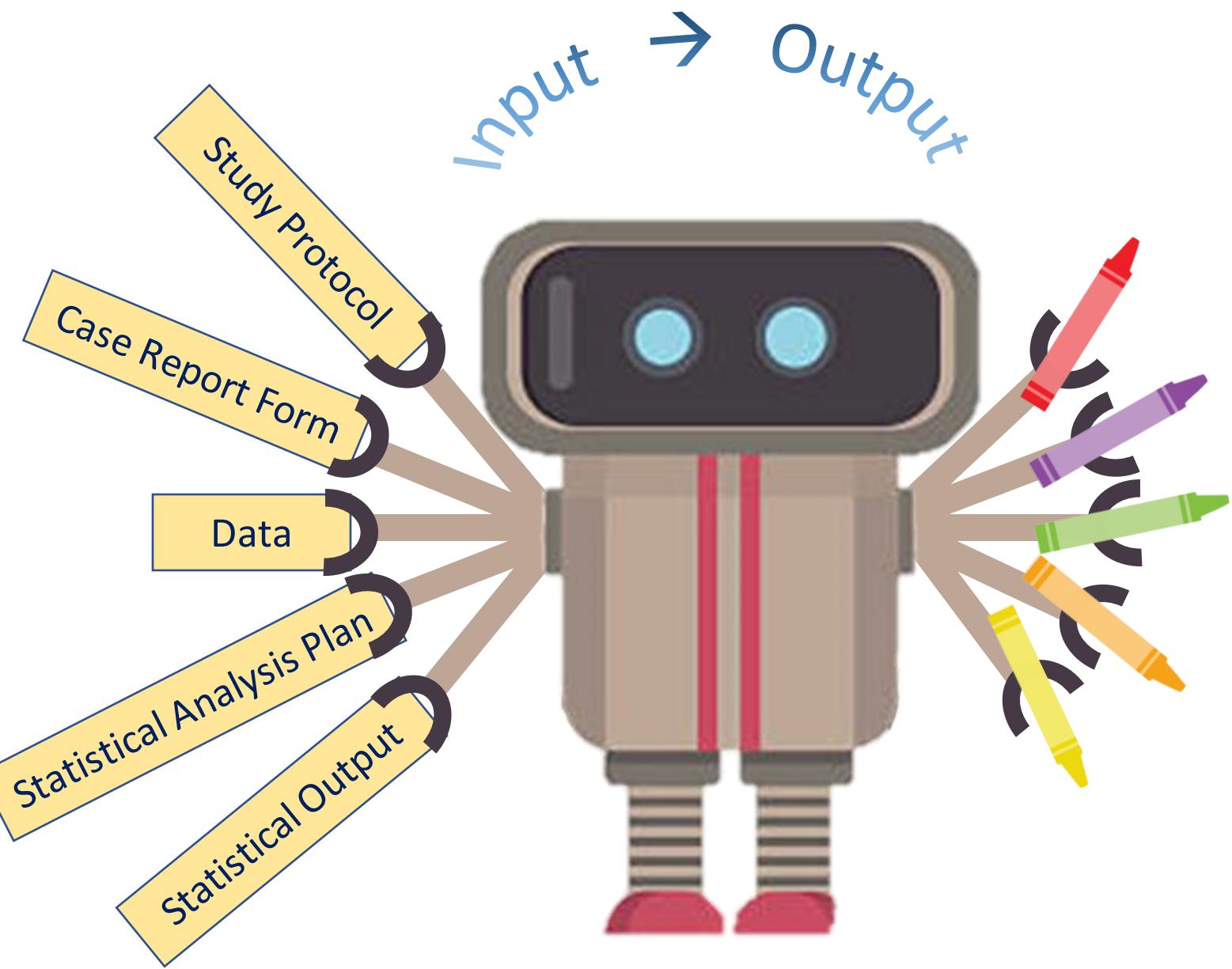
Implementation



Next Steps



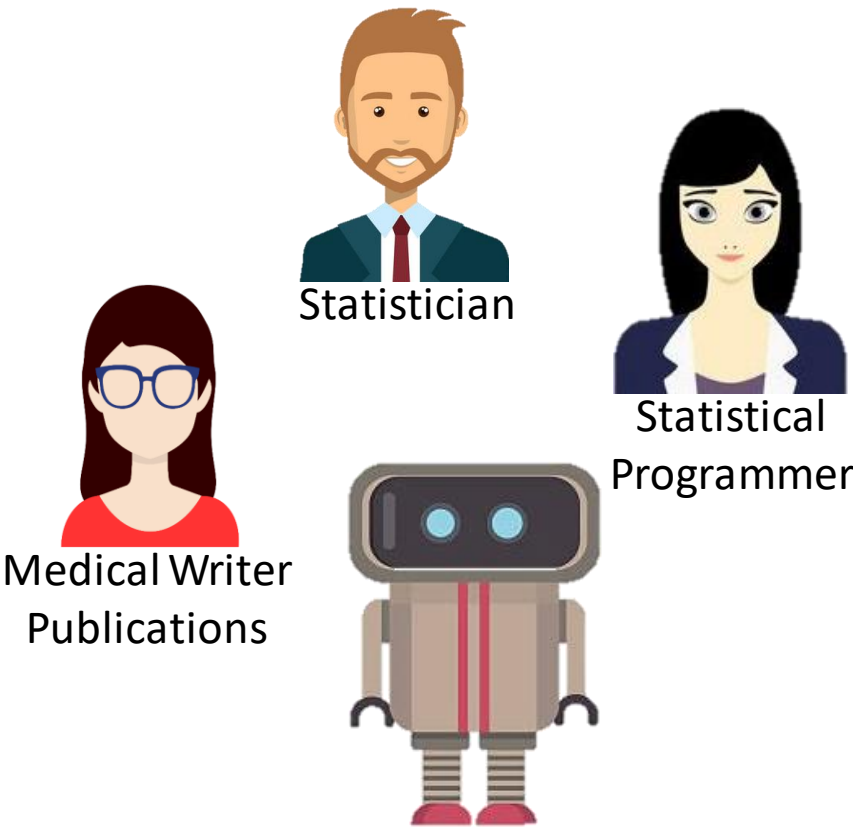
Design: Structure of Automated Text



SPECDOC Components:

	1	Study	• Description of study
	2	Data	• Description of the data • Database lock/interim cut • Endpoint • Missing data handling
	3	Analysis Details	• Full specification of the statistical method
	4	Results	• Comprehensive and consistent description of the results
	5	Statistical Methods Primer	• Concept, video tutorial and other resources on statistical methods for reference

Design: Templates



Example Template for ANCOVA:

Separate analysis of covariance (ANCOVA) models were implemented using the <PROC> procedure for each of the following endpoint(s) and visits:

Endpoint(s)	Response variable	Timepoints (excluding baseline)
PARAM1	CHG or PCHG or AVAL, etc.	Visits for PARAM1
PARAM2	CHG or PCHG or AVAL, etc.	Visits for PARAM2
...	...	...

For each of the endpoint(s) shown above, <_CATCOVLB> as categorical covariate(s), <_CONTCOVLB> as continuous covariate(s), and <_ADDINTERXLB> as interaction terms were used.

<_RANDEFFECTLB> was considered a random effect on the intercept with a covariance type of Variance Components (VC); that is, the intercept is allowed to be different for each <_RANDEFFECTLB> and the measurements within each <_RANDEFFECTLB> are considered independent.

At baseline (Analysis <AWU> = <AW target at the baseline visit>), only unadjusted descriptive statistics were displayed for each level of the categorical covariate(s).

At post-baseline visits, the unadjusted descriptive statistics, LS-means with two-sided <1-alpha level>% confidence intervals, p-value to test whether the LS-means of <change from baseline, or percent change from baseline> to a given visit of a particular level of a categorical variable (not the difference between levels) were different from 0. The differences in LS-means with two-sided <1-alpha level>% confidence intervals and p-value were also computed for each categorical covariate relative to its reference level after adjusting for the aforementioned covariates.

For each categorical covariate, the following reference groups were used:

Covariate	Reference	Other covariate values
Covariate 1	Reference 1	Non-reference values for Covariate 1
Covariate 2	Reference 2	Non-reference values for Covariate 2
...	...	...

Hedge's G and Cohen's D to measure the effect size are calculated.

Overview



Introduction



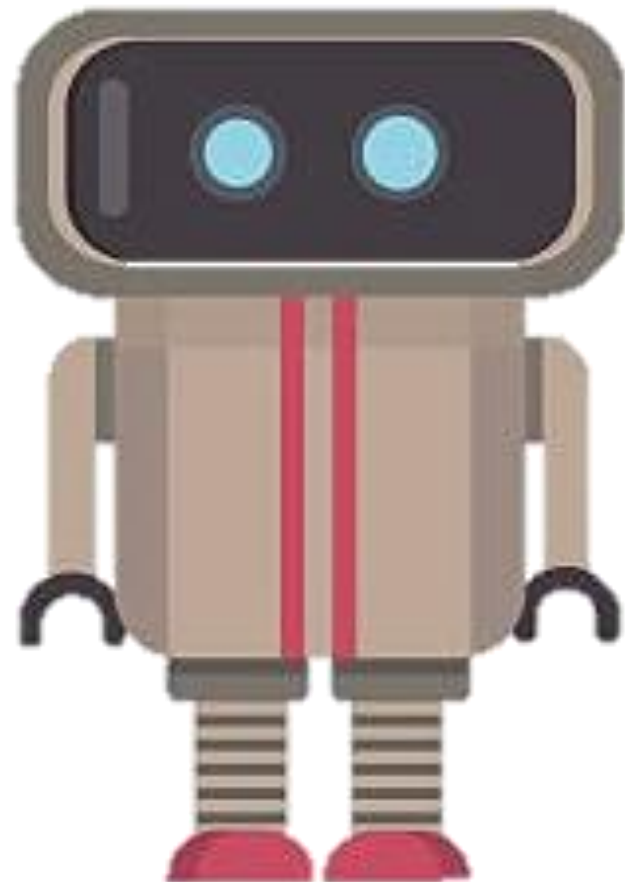
Design



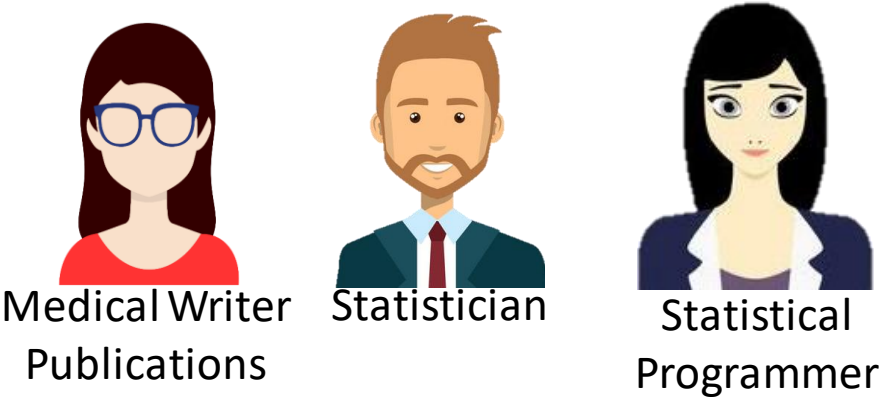
Implementation



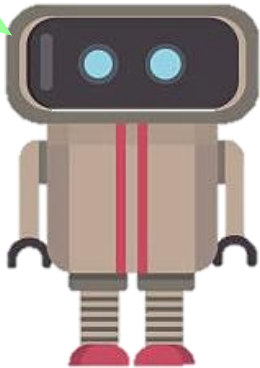
Next Steps



Implementation



Move over humans. I got this...



Example Template for ANCOVA:

Separate analysis of covariance (ANCOVA) models were implemented using the `<PROC>` procedure for each of the following endpoint(s) and visits:

Endpoint(s)	Response variable	Timepoints (excluding baseline)
PARAM1	CHG or PCHG or AVAL, etc.	Visits for PARAM1
PARAM2	CHG or PCHG or AVAL, etc.	Visits for PARAM2
...	...	...

For each of the endpoint(s) shown above, `<_CATCOVLB>` as categorical covariate(s), `<_CONTCOVLB>` as continuous covariate(s), and `<_ADDINTERXLB>` as interaction terms were used.

`<_RANDEFFECTLB>` was considered a random effect on the intercept with a covariance type of Variance Components (VC); that is, the intercept is allowed to be different for each `<_RANDEFFECTLB>` and the measurements within each `<_RANDEFFECTLB>` are considered independent.

At baseline (Analysis `<AWU>` = `<AW target at the baseline visit>`), only unadjusted descriptive statistics were displayed for each level of the categorical covariate(s).

At post-baseline visits, the unadjusted descriptive statistics, LS-means with two-sided `<1-alpha level>`% confidence intervals, p-value to test whether the LS-means of `<change from baseline, or percent change from baseline>` to a given visit of a particular level of a categorical variable (not the difference between levels) were different from 0. The differences in LS-means with two-sided `<1-alpha level>`% confidence intervals and p-value were also computed for each categorical covariate relative to its reference level after adjusting for the aforementioned covariates.

For each categorical covariate, the following reference groups were used:

Covariate	Reference	Other covariate values
Covariate 1	Reference 1	Non-reference values for Covariate 1
Covariate 2	Reference 2	Non-reference values for Covariate 2
...	...	...

Hedge's G and Cohen's D to measure the effect size are calculated.

Implementation

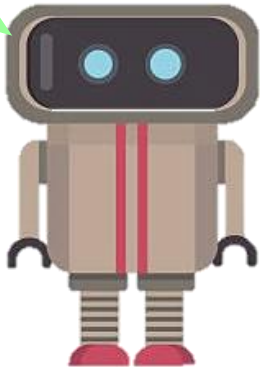
Example Template for ANCOVA:

Separate analysis of covariance (ANCOVA) models were implemented using the MIXED procedure for each of the following endpoint(s) and visits:

Endpoint(s)	Response variable	Timepoints (excluding baseline)
Disease Activity Index	Value at Visit (AVAL)	Week 4, Week 8, Week 12, Week 24

For each of the endpoint(s) shown above, Treatment (ARM), Region (STRATUM), Previous Medications for the Indication (STRATUM2) as categorical covariate(s) and Baseline Disease Activity Index (BASE) as continuous covariate(s) were used.

Move over humans. I got this...



At baseline (Analysis week = 0), only unadjusted descriptive statistics were displayed for each level of the categorical covariate(s).

At post-baseline visits, the unadjusted descriptive statistics, LS-means with two-sided 95% confidence intervals. The differences in LS-means with two-sided 95% confidence intervals and p-value were also computed for each categorical covariate relative to its reference level after adjusting for the aforementioned covariates.

For each categorical covariate, the following reference groups were used:

Covariate	Reference	Other covariate values
Treatment	Placebo	Treatment Dose 1, Treatment Dose 2
Region	Rest of World	Western, South America
Previous Medications for the Indication	1	>1

Hedge's G and Cohen's D to measure the effect size are calculated.

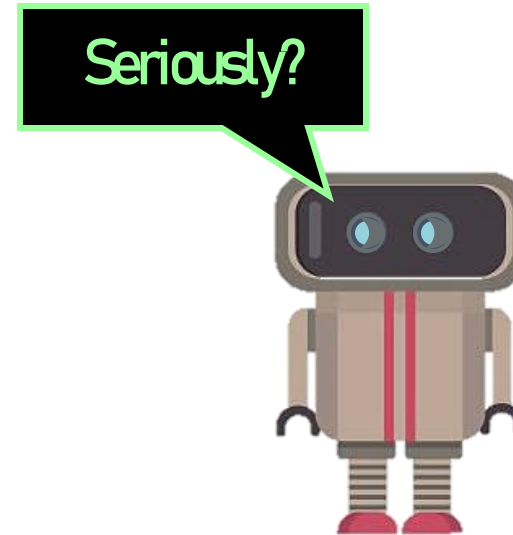
Study and Data Description Sections

Reports often start with a high-level summary of the clinical study as well as a description of the dataset used to generate the analysis



Detective Work

Medical writers search the protocol, statistical analysis plan, and other documents manually



SPECBOT

SPECBOT uses centrally stored metadata

Study and Data Description Metadata: TS and ADTS

- CDISC Trial Summary (TS) domain from SDTM contains some of the information
- Following CDISC guidelines and principles, ADTS was created to store additional information
- Below are select examples of TS and ADTS used to populate the Study and Data Description sections

TS Domain

Trial Summary Parameter Test Code (TSPARMCD)	CDISC SDTM Trial Summary Parameter Short Name Terminology
DCUTDESC	Data Cutoff Date Description
DCUTDTC	Data Cutoff Date
HLTSUBJI	Healthy Subject Indicator
THERAREA	Therapeutic Area
INDIC	Trial Indication
LENGTH	Trial Length
TBLIND	Trial Blinding Schema
TPHASE	Trial Phase
...	

ADTS

Trial Summary Parameter Test Code (TSPARMCD)	Trial Summary Parameter Short Name Terminology
STUDYNM	Study name
DCUTPOSE	Data Cutoff Purpose
DCUTSPE	Data Cutoff Prespecify Description
SASVERS	SAS Version
EPTNM	Endpoint Name
EPTNMLEV	Endpoint Level
ANLPRESP	Endpoint Prespecified Description
...	

Implementation

Specdocs

Specdoc for Table 2.2.1

Analysis Description

Analysis Result Metadata (ARM)

Results Description

Primer of Methods

	1	Study	Description of study
	2	Data	- Description of the data - Database lock/interim cut - Endpoint - Missing data handling
	3	Analysis Details	Full specification of the statistical method
	4	Results	Comprehensive and consistent description of the results
	5	Statistical Methods Primer	Concept, video tutorial and other resources on statistical methods for reference

Study	These are outputs from the [REDACTED] PHASE III TRIAL study [REDACTED] in Adult Patients with [REDACTED]
Datacut	This is a Final Database Lock cutoff prespecified in Protocol with a cutoff date of [REDACTED]. The objective of this cutoff was: Clinical Study Report (CSR). For this data cut, the extent of data cleaning was: 100%. The cutoff was unblinded. The following groups were unblinded for this cutoff: study team.
Analysis	These analyses were Pre-specified in protocol. The following endpoints were included in this analysis: [REDACTED] Secondary Endpoint). The following visits were included in this analysis: Baseline, Week [REDACTED]. The Intent-to-Treatment (ITTFL) included Patients who were randomized. Missing data were imputed as follows: Observed cases, censored after rescue use. SAS version number is 9.4.
Missing data handling	Missing data were first imputed as follows: Observed Cases. The details are as follows: Observed cases, censored after rescue use.
ANCOVA Documentation	Separate analysis of covariance (ANCOVA) models were implemented using the MIXED procedure for each of the following endpoint(s) and visits:

Endpoint(s)	Response variable	Timepoints (excluding baseline)
[REDACTED]	Value at Visit (AVAL)	Week [REDACTED]

For each of the endpoint(s) shown above, Treatment (ARM), Region (STRATUM), Previous [REDACTED] as categorical covariate(s), Baseline [REDACTED] (BASE) as continuous covariate(s) were used.

At baseline (Analysis week = 0), only unadjusted descriptive statistics were displayed for each baseline of the following baseline covariates:



At baseline, for the categorical covariate of Treatment, there were ██████████ and ██████████ subjects in Placebo (reference), ██████████ and ██████████ respectively. The observed means (Standard Error, or SE) for these 3 groups were ██████████ and ██████████ respectively. The ██████████ category had the lowest observed mean, and the placebo category had the highest observed mean, numerically.

At Week ██████████ for the categorical covariate of Treatment, there were ██████████ and ██████████ subjects in Placebo (reference), ██████████ and ██████████ respectively. The observed means (Standard Error, or SE) for these 3 groups were ██████████ and ██████████ respectively. The covariate-adjusted LS means (95% CI) from the ANCOVA model for these 3 groups were ██████████ and ██████████ respectively.

For the categorical covariate of Treatment, the LS mean differences (95% CI) of ██████████ were: ██████████ with a p-value of ██████████ for ██████████ vs placebo and ██████████ with a p-value of ██████████ for ██████████ vs placebo.

The effect size vs. reference (95% CI) ██████████ vs. placebo, at ██████████ vs. placebo.

ANCOVA
ANCOVA, or ANalysis of COVariance, models continuous outcome (y) vs. a collection of continuous and/or categorical covariate. In the context of evaluating a new treatment, this can be interpreted as modeling the treatment effect (usually as one of the categorical covariates), while controlling for other continuous and/or categorical factors (e.g., region and baseline disease status).

P-value
A p-value is a measure of the probability that an observed difference could have occurred just by random chance. Specifically, the p-value is the probability of obtaining results at least as extreme as the observed results relative to the null hypothesis in a statistical hypothesis test, assuming that the null hypothesis is correct. (See MA Mini Learning on Demand, p-value). When a p-value is smaller than 0.05, the corresponding 95% confidence interval excludes the null hypothesis (a null hypothesis is no treatment effect, or 0 in a superiority comparison).

Overview



Introduction



Design



Implementation



Next Steps



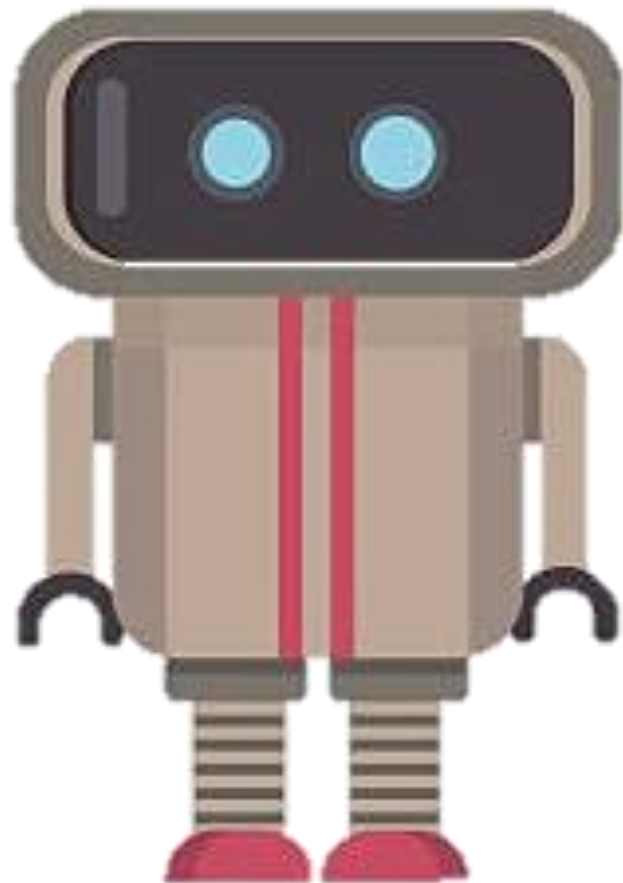
Next Steps



ChatGPT



OpenAI



Conclusion



Consistency



Comprehensive



Speed



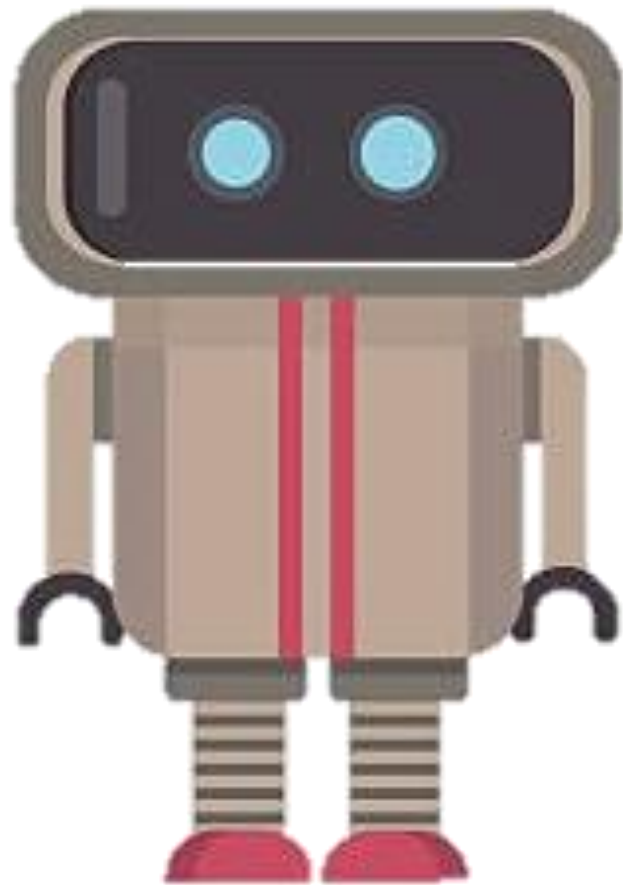
Efficiency



Accuracy



Transparency



Do any of you humans
have any QUESTIONS?