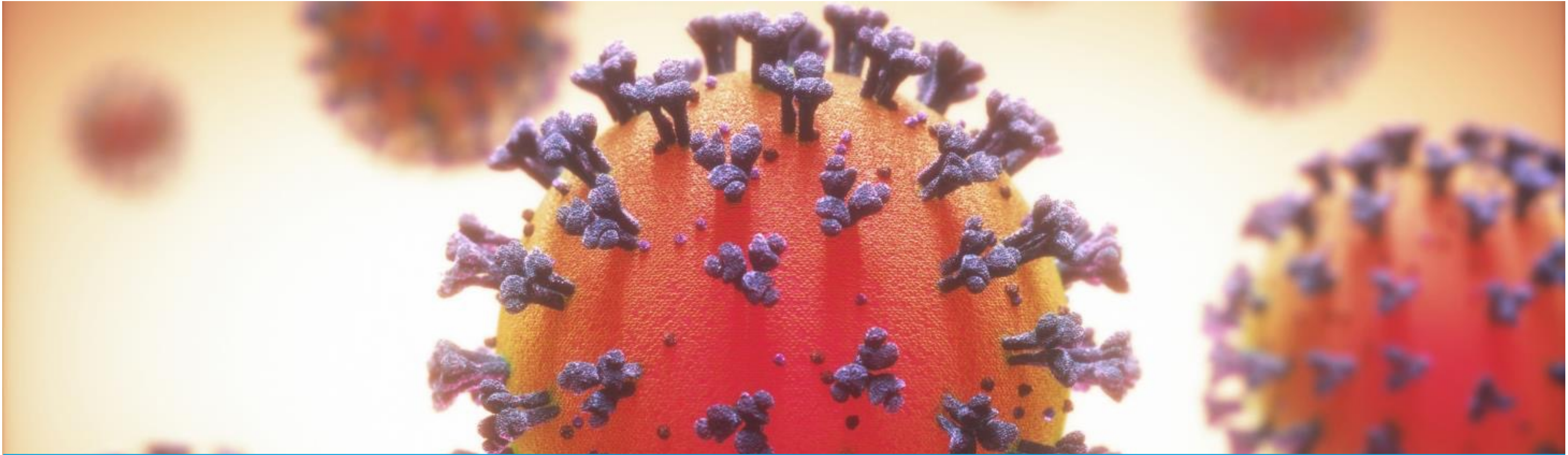




Harnessing the Power of Shiny to Visualize Missing Data Due to the COVID-19 Pandemic

Sangeeta Bhattacharya and Mark Rothe

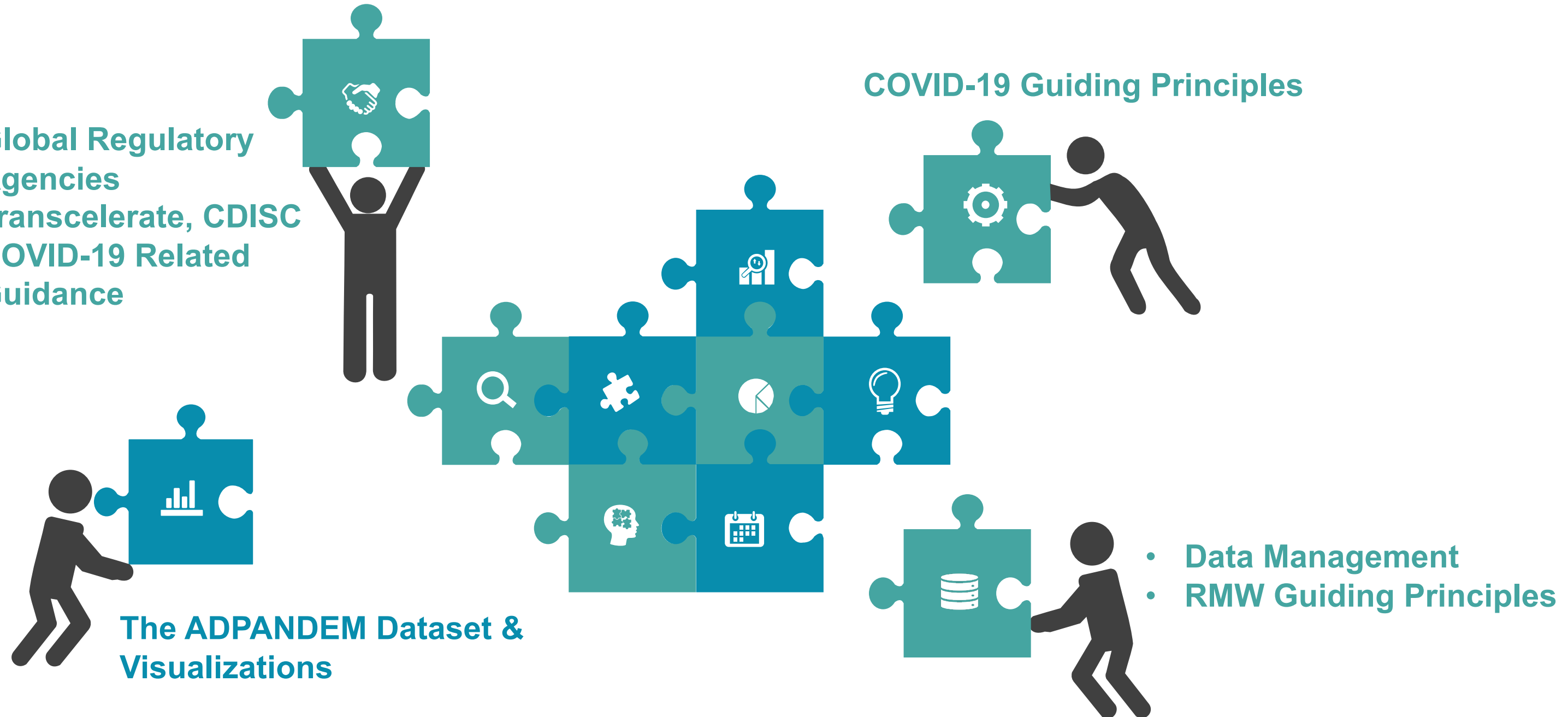
14-June-2021

- sbhatt11@its.jnj.com
- mrothe@its.jnj.com

Overcoming the Pandemic: *The “ADPANDEM”* Story

- Global Regulatory Agencies
- Transcelerate, CDISC
- COVID-19 Related Guidance

COVID-19 Guiding Principles



Standard Reporting Guidelines

The **Guiding Principles for Standard Reporting** addresses COVID-19 impact to Displays (Tables, Listing, Graphs) related to:

- **Protocol Deviations**
- **Treatment / Exposure**
- **Subject Disposition & Discontinuation**
- **Efficacy Estimands / End Points**
- **Visits**
- **Adverse Events & Deaths**
- **Laboratory Data**
- **Concomitant Medications**

Importantly, the Guiding Principles also include a template supplemental **Data Presentation Specification (DPS)** that outlines the Displays that study teams need to consider. Each TA can use the standard display or modify as needed.

Related Guidance: COVID-19 Protocol Deviations Guiding Principles; COVID-19 Data Collection Guidance; Implications of COVID-19 on Statistical Aspects; COVID-19 AE Coding Guidance

Birth of the Analysis Dataset “ADPANDEM”

To aid in the assessment of COVID-19 impact, the creation of an ADaM compliant pandemic analysis dataset—**ADPANDEM**—was created.



- The dataset contains records for all possible visits (done or missing) per subject for specific important endpoints (parameters).
- There are flags for assessments that are assumed “missed due to COVID-19”. (*Logic takes into account missed assessments and COVID-19 related PDs & AEs for the patient.*)
- Search criteria for COVID-19 AEs is updated periodically based on upgraded MedDRA versions.
- Both Major and minor protocol deviations due to COVID-19 are mapped to DV.
- COVID-19 PDs are prefixed with “COVID-19”.

Note: The logic and input data in the creation of ADPANDEM may be modified based on the best / most reliable source of information for a given study—especially, if the required data is collected via the introduction of CRFs versus alternate sources.

ADPANDEM Details

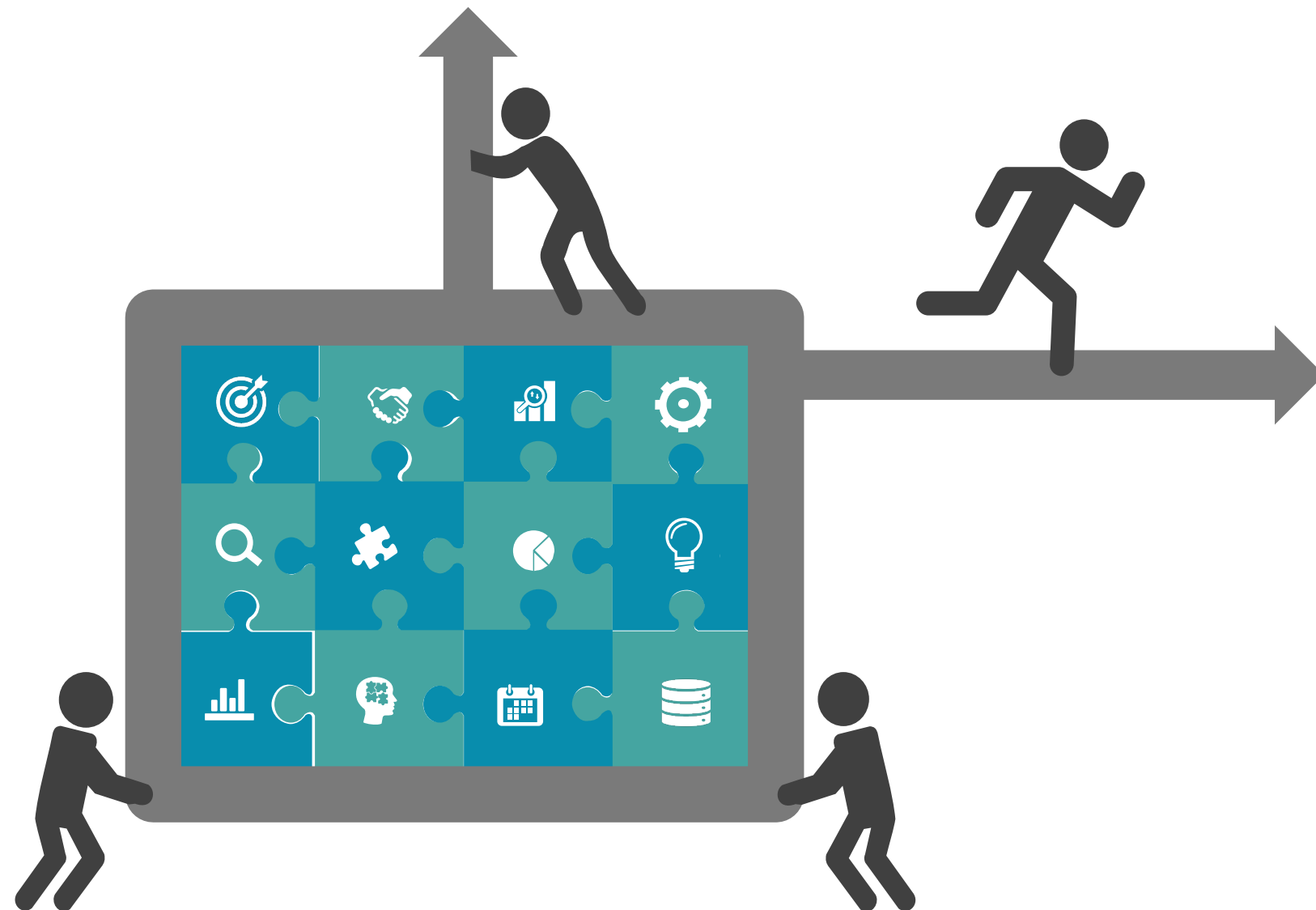
TSIDS-ST06-COV:Study Assessment Compliance; [Analysis Set] (Study [Study ID])					
	[Active Study Agent]				
	[Placebo]	[Treatment Group A]	[Treatment Group B]	Combined	Total
Analysis set: [Randomized]	###	###	###	###	###
Median follow-up time ([months])	xx.x ####	xx.x ####	xx.x ####	xx.x ####	xx.x ####
% of subjects who missed at least one scheduled [study assessment/disease evaluation]	### (xx.x%) ####	### (xx.x%) ####	### (xx.x%) ####	### (xx.x%) ####	### (xx.x%) ####
% of subjects who missed at least one scheduled [study assessment/disease evaluation] due to COVID-19	### (xx.x%) ####	### (xx.x%) ####	### (xx.x%) ####	### (xx.x%) ####	### (xx.x%) ####
% of subjects who missed 2 or more consecutive scheduled [study assessment/ disease evaluation]	### (xx.x%) ####	### (xx.x%) ####	### (xx.x%) ####	### (xx.x%) ####	### (xx.x%) ####
% of subjects who missed 2 or more consecutive scheduled [study assessment/ disease evaluation] due to COVID-19	### (xx.x%) ####	### (xx.x%) ####	### (xx.x%) ####	### (xx.x%) ####	### (xx.x%) ####
# of missed [study assessment/ disease evaluation] per subject					
Mean (SD)	xx.x (xx.xx) ####	xx.x (xx.xx) ####	xx.x (xx.xx) ####	xx.x (xx.xx) ####	xx.x (xx.xx) ####
Median	xx.x ####	xx.x ####	xx.x ####	xx.x ####	xx.x ####
Range	(xx.xx) ####	(xx.xx) ####	(xx.xx) ####	(xx.xx) ####	(xx.xx) ####
IQ range	(xx.xx) ####	(xx.xx) ####	(xx.xx) ####	(xx.xx) ####	(xx.xx) ####
# of missed [study assessment/ disease evaluation] due to COVID-19 per subject					
Mean (SD)	xx.x (xx.xx) ####	xx.x (xx.xx) ####	xx.x (xx.xx) ####	xx.x (xx.xx) ####	xx.x (xx.xx) ####
Median	xx.x ####	xx.x ####	xx.x ####	xx.x ####	xx.x ####
Range	(xx.xx) ####	(xx.xx) ####	(xx.xx) ####	(xx.xx) ####	(xx.xx) ####
IQ range	(xx.xx) ####	(xx.xx) ####	(xx.xx) ####	(xx.xx) ####	(xx.xx) ####

Supports the creation of the table (on the left)—an example from the list of suggested tables.

- Padded dataset: 1 record per subject, per visit, per assessment / endpoint.
- Records are inserted for actual and estimated visits through study discontinuation date or today's date.
- Flags those visits that are missed for easy population of tables.

Use of ADPANDEM

Analysis-assessing the impact of missing end points on trial validity.



Monitoring-identifying and mitigating missing data by carefully planning and implementing strategies to minimize the likelihood of missing data.

Visualizing COVID-19 Missingness



How best to visualize missingness due to COVID-19?

- ADPANDEM keys: **STUDYID**, **USUBJID**, **AVISITN**, and **PARAMCD**.
- Multiple visits and parameters per subject.
- Display missingness by country, site, visit, and parameter.

Solution:

- The R package **shiny**
- **Cumulative** view of missingness over time.
- **By-visit** view of missingness.

Coding aside: How to decode country names from ISO 3166-1 country codes?
→ Scrape Wikipedia via the R package **rvest**.

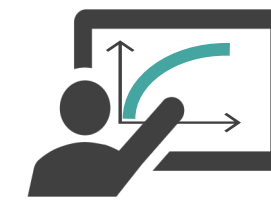
Cumulative View of Missingness: Summary



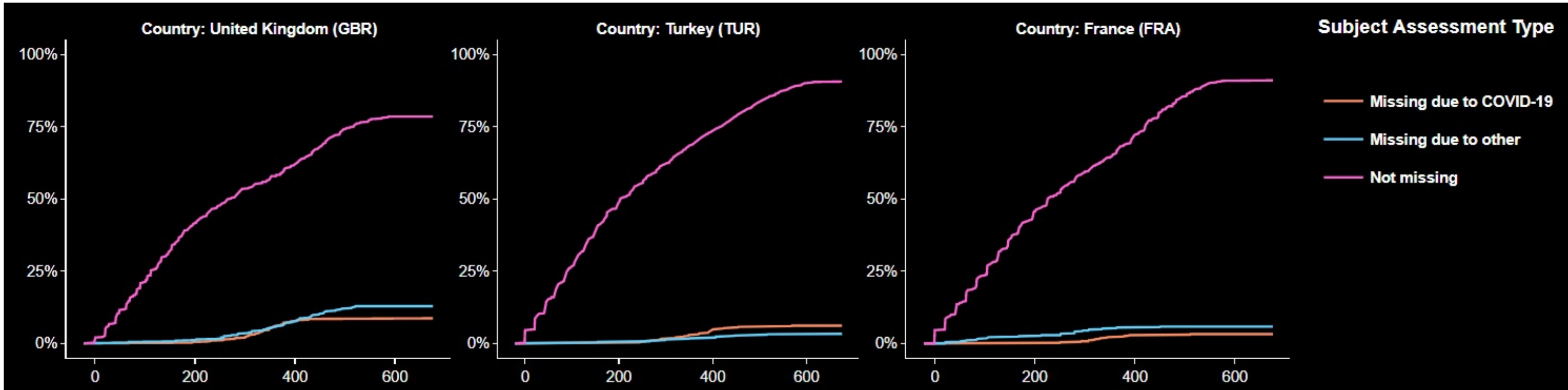
- Display **red-amber-green** (RAG) heatmap of cumulative COVID-19 missingness by **parameter**, **country**, and **site**:
- **n** : Cumulative (total across all study days) number of subject assessments missed due to COVID-19.
 - **N** : Cumulative (total across all study days) number of subject assessments.
 - **%** : $n / N \times 100$

↕ Parameter Metadata	↕ Country	↕ % (n / N)	↕ n	↕ N
• Chemistry (MCLP) (Code: MCLPA — Parameter: Monoclonal protein - Assessment done — Source: SDTM.LB)	Brazil (BRA)	1.1%	14	1275
	Canada (CAN)	2.64%	23	870
	↕ ↔ Site ID	↕ % (n / N)	↕ n	↕ N
	CA10002	6.95%	21	302
	CA10003	2.08%	1	48
	CA10006	0.8%	1	125

Cumulative View of Missingness: Plots



→ Plot cumulative COVID-19 % missingness by study day and facet by permutations of **parameter**, **country**, and **site**.

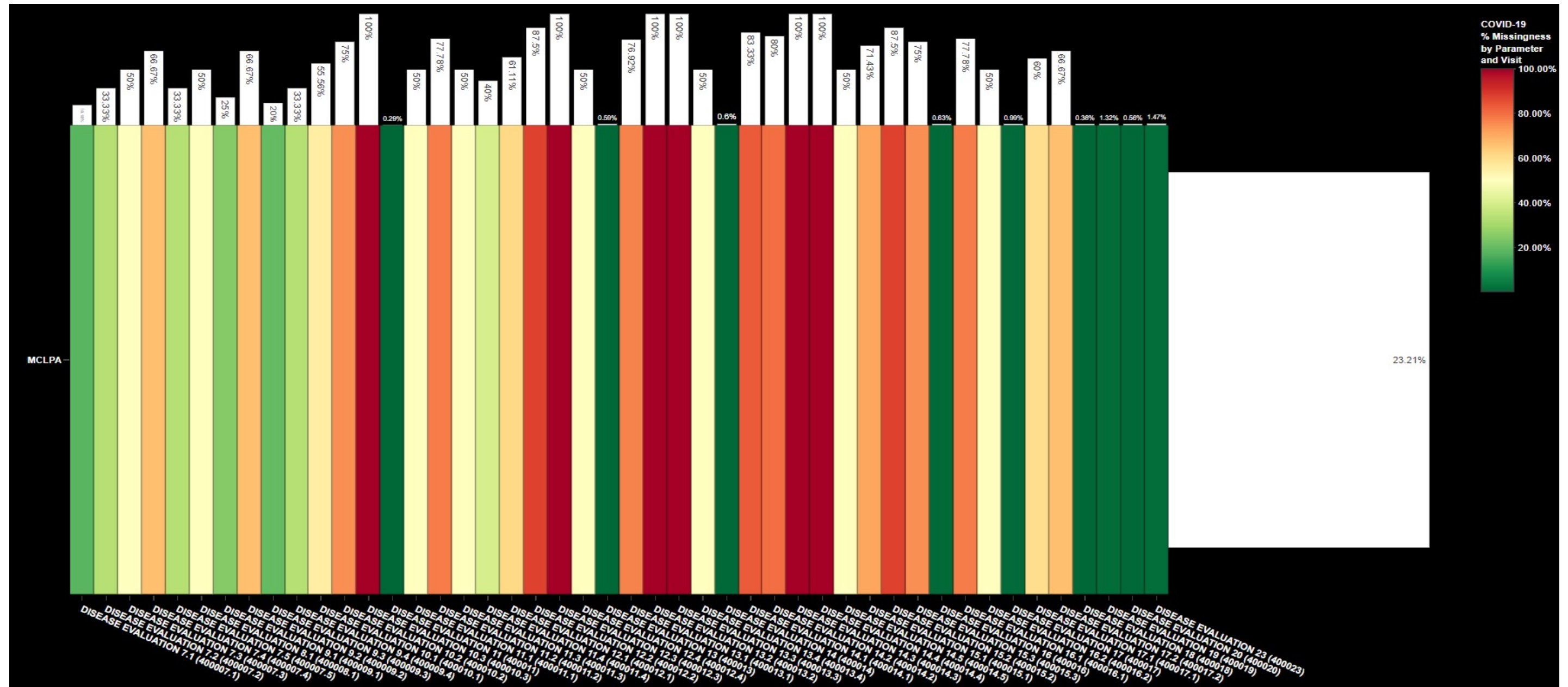


Coding aside: Faceted via the R package ggplot2. In hindsight, the **trelliscopejs** R package would have afforded greater flexibility without ggplot2's screen real estate challenges.

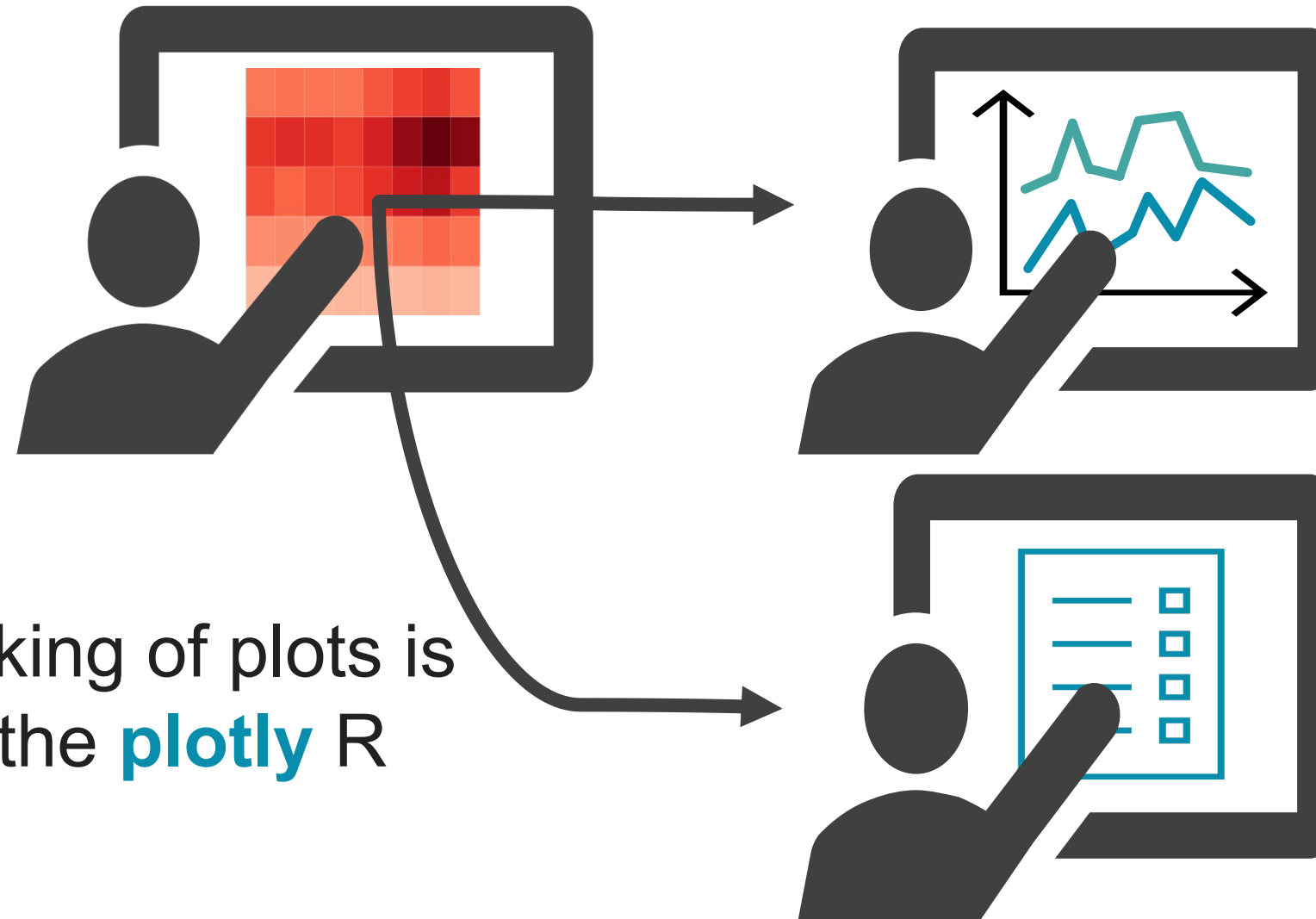
By-Visit View of Missingness: Heatmap



→ Create heatmap of COVID-19 % missingness by parameter and visit.

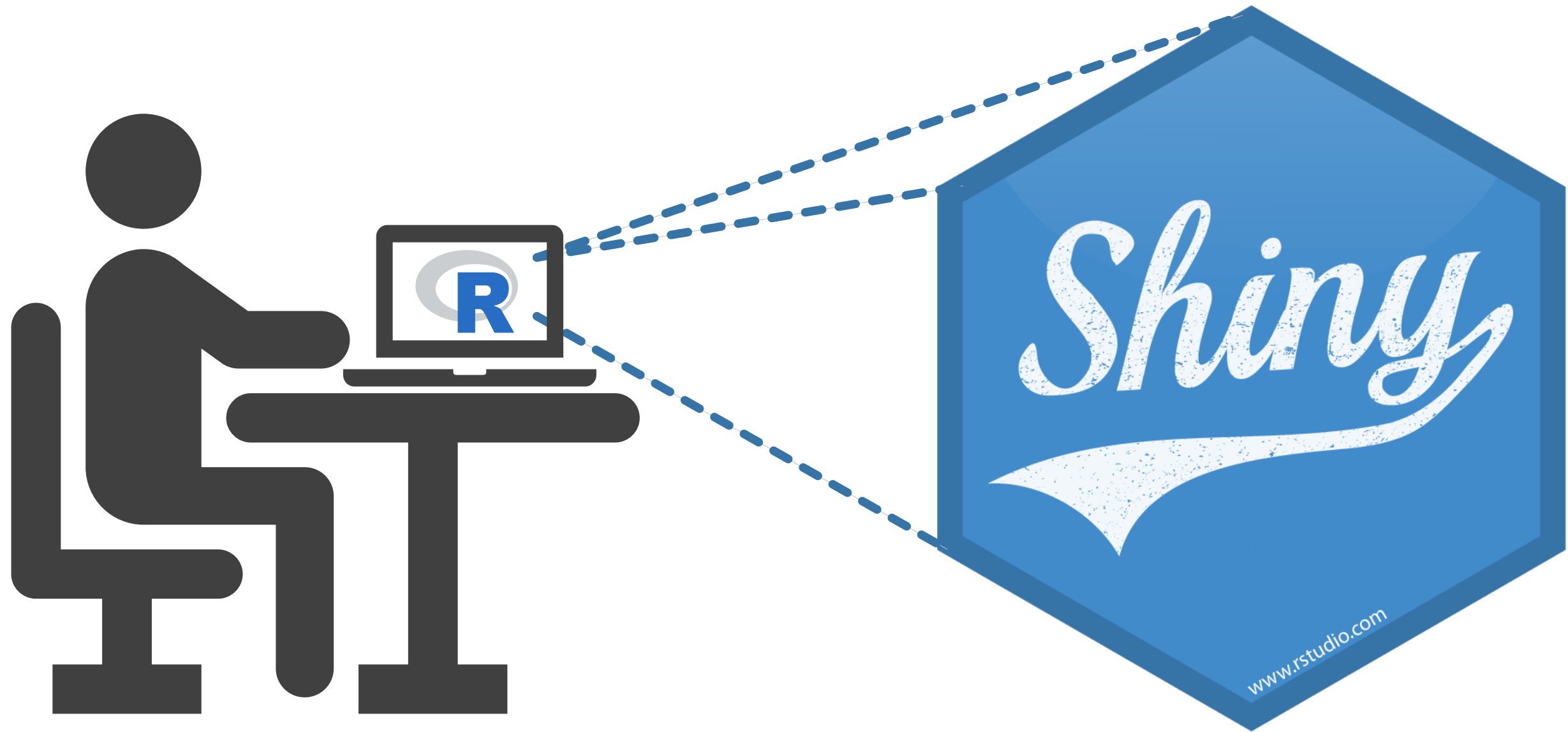


By-Visit View of Missingness: Drill Down



Coding aside: Linking of plots is made possible via the **plotly** R package.

Live Demo of the ADPANDEM R Shiny App



Future ADPANDEM Developments



Additional COVID-19 guidance is forthcoming, necessitating updates to ADPANDEM and interactive data visualizations ...

- FDA's **Office of Oncologic Diseases (OOD)** Divisions and Programs is interested in missingness as it relates to the type of patient contact (**remote** vs. **onsite**).
- **MHRA** wants to investigate the impact of COVID-19 vaccination on clinical trials.

