

# Workstream 1

## Retrospective Data Analysis

May 4, 2011



# Overview

- Background
- Data Cleaning
- Results
- Limitations
- Lessons Learned

# Background

- Data received from 18 organizations:
  - Academic (2)
  - CRO-ARO (4)
  - Biotechnology (2)
  - Device (2)
  - Government (1)
  - Investigators (2)
  - Pharmaceutical (6)
- Objective was to collect “simple” metrics

# Retrospective Collection - Data Elements

Field Name	Suggested Definition
Site ID	List unique site identifier as per your system
Protocol ID	List a protocol identifier as per your system
Site Type	Type of site conducting the study (Academic, Hospital based, private practice, independent, VA)
IRB Type	Central, Local, Regional
Therapeutic Area	CV/Metabolic, Hematology & Oncology, Infectious Diseases, Neurosciences, Vaccines, Ophthalmology, Pain, Dermatology, Inflammation, Allergy & Respiratory, Urology, Women's Health, Device and Other
If TA is "Other" - Specify	If TA is other than one of the options given; enter specific TA
Sponsor Name	List name of the sponsor for the study

Field Name	Suggested Definition
Date protocol sent by sponsor/CRO to site	Date the protocol was sent by the sponsor/CRO to the site. If as a site you do not know this date, leave this field blank.
Date protocol received by site	Date the protocol was received by the site. If as a sponsor/CRO you do not know this date, leave this field blank.
Date of protocol submitted by site for IRB approval	Date the protocol is submitted by the site to the IRB of record
Date of final IRB decision on the protocol	Date the IRB approves the protocol
Date final signature obtained to fully execute site contract	Date the final signature is obtained on the contract and contract is considered fully executed
Date site enrolled first patient	Date the site enrolled the first patient into the study (date informed consent signed)

# Data Cleaning

Evaluated blank fields to determine if it could be reasonably derived from other fields provided in the sample.

Examples:

- Identification of IRB type from name of IRB
- Identification of site type from name of site
- Reassigning therapeutic area to a broader category
- Filtered a site's data to include uniform phase data; i.e., all from Phase II, Phase III, or Device studies
- Some data submitted included calculated cycle times without source dates; data still used when we could identify a match with our cycle time definitions

# Data Cleaning

- For cycle times, did not include any cycle times that were negative
- Different cutpoints were considered, but all seemed arbitrary
- Cutpoint of zero allowed outlier observations to still be included in the analysis

# Data Cleaning

Organization Type	# Original Data Lines	# Data Lines After Cleaning
Academic	4,516	559
ARO-CRO	703	703
Biotech	512	512
Device	79	79
Govt	51	51
Investigators	43	43
Pharma	4,769	3,449
<b>TOTAL</b>	<b>10,673</b>	<b>5,396</b>

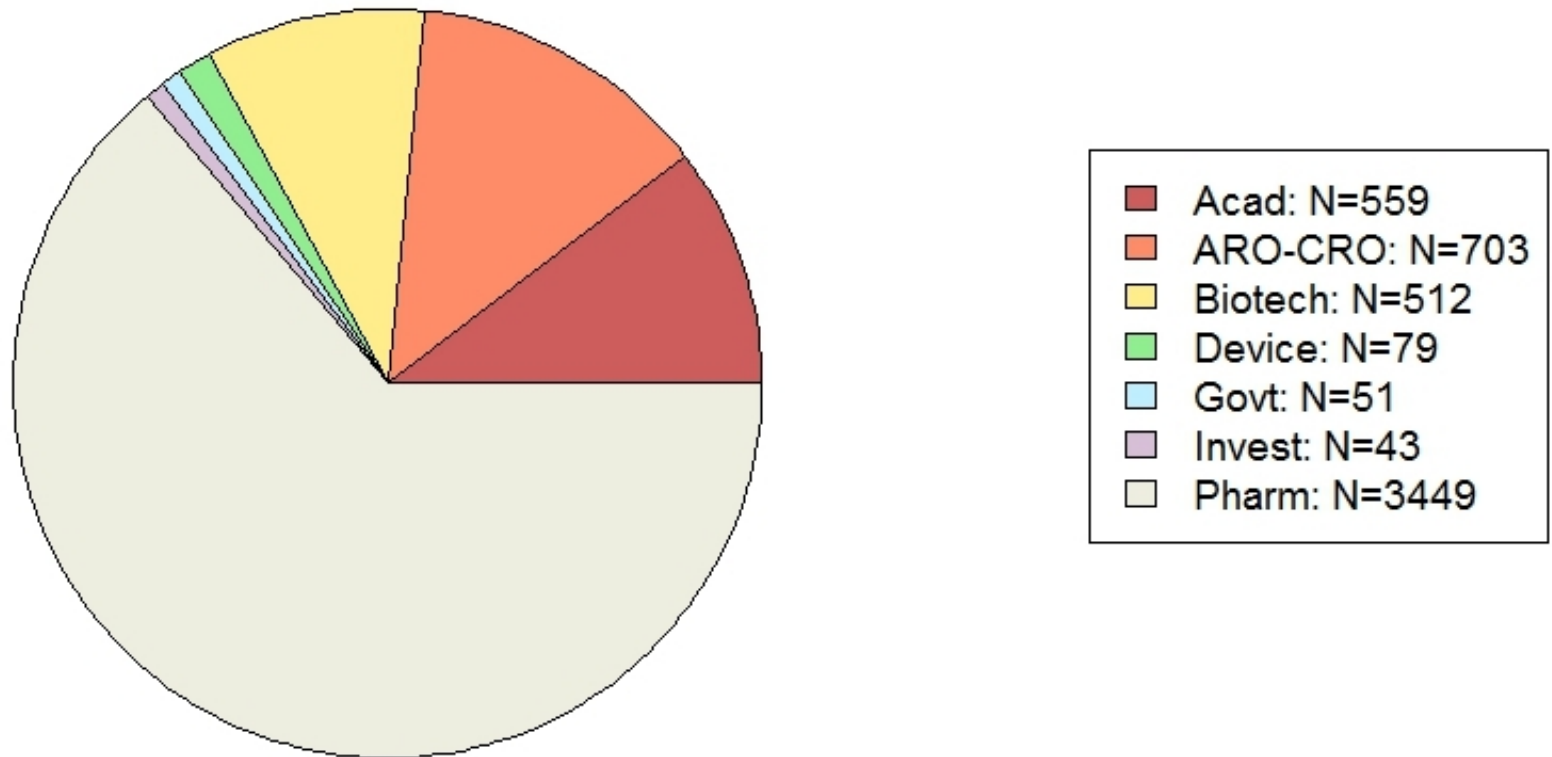
# Results Overview

- Overview of characteristics of data submitted
  - ◆ Submitting organization
  - ◆ Site type
  - ◆ IRB type
  - ◆ Therapeutic area
  
- Missing Data
  
- Cycle time metrics with various stratifications
  - ◆ Protocol sent/received to protocol submitted to IRB
  - ◆ Protocol sent/received to IRB final decision
  - ◆ Protocol sent/received to contract execution
  - ◆ Protocol sent/received to 1<sup>st</sup> patient enrolled
  - ◆ Protocol submission to IRB to IRB final decision
  - ◆ IRB final decision to 1<sup>st</sup> patient enrolled
  - ◆ Contract executed to 1<sup>st</sup> patient enrolled

Who Submitted the Data?  
How Much Data Were Received?

# Results

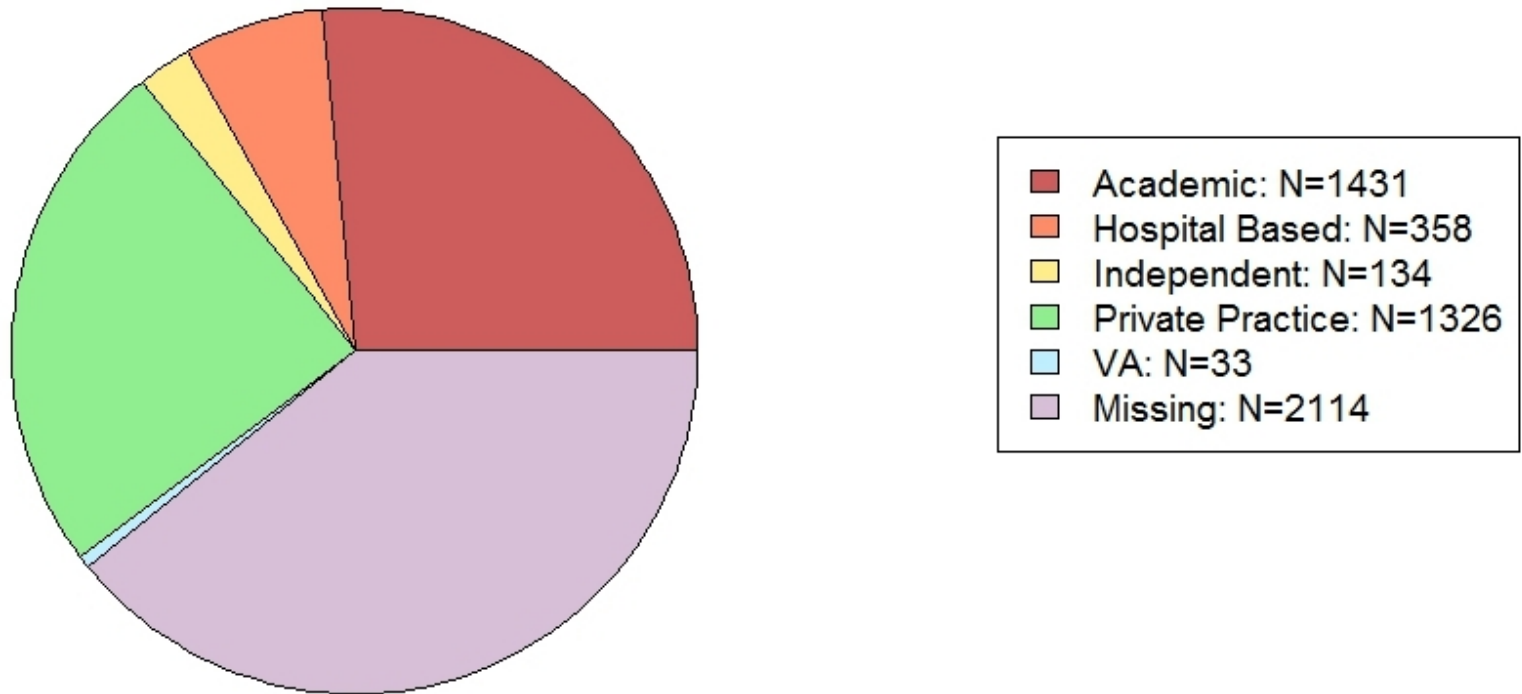
Totals by Organization Type



# What is the Distribution of Site Types in the Data?

# Results

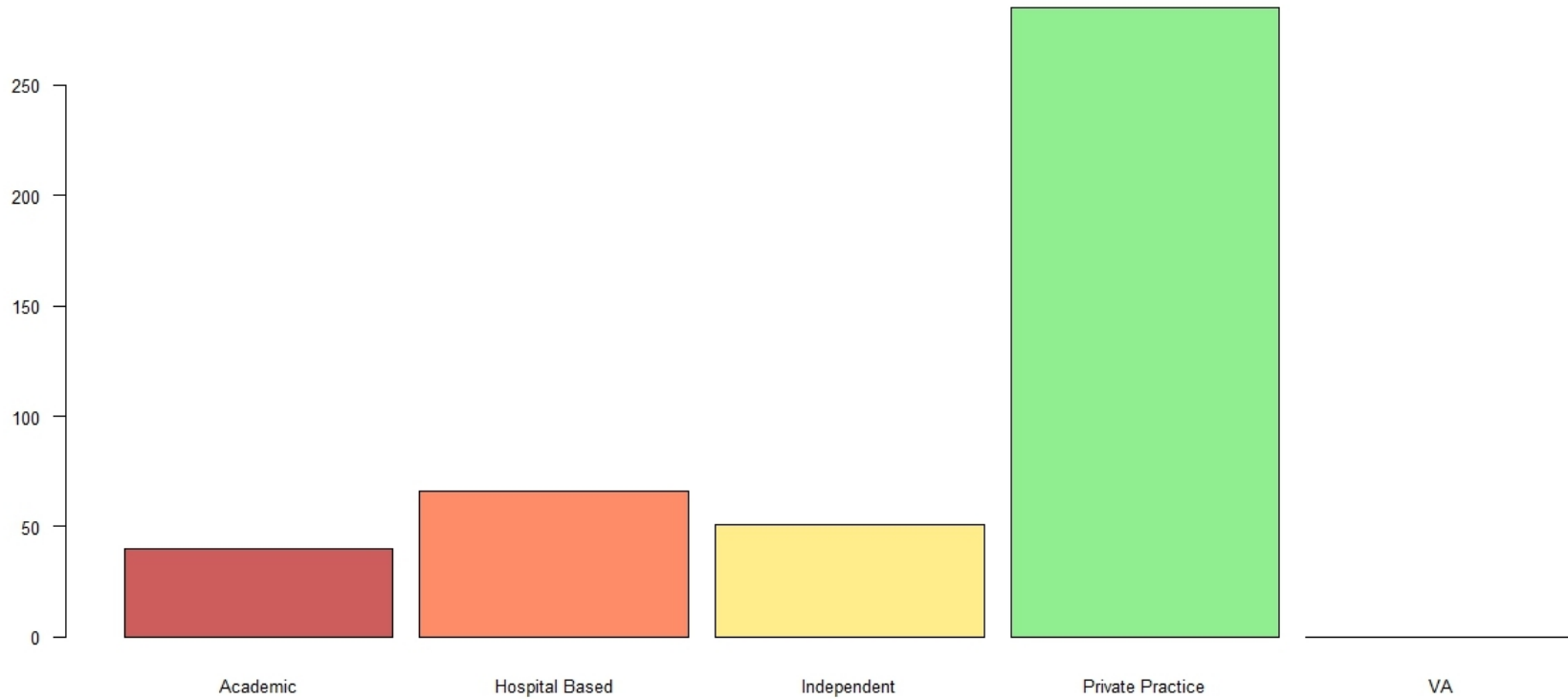
Totals by Site Type



Within Each Organization,  
What is the Distribution  
of Site Type?

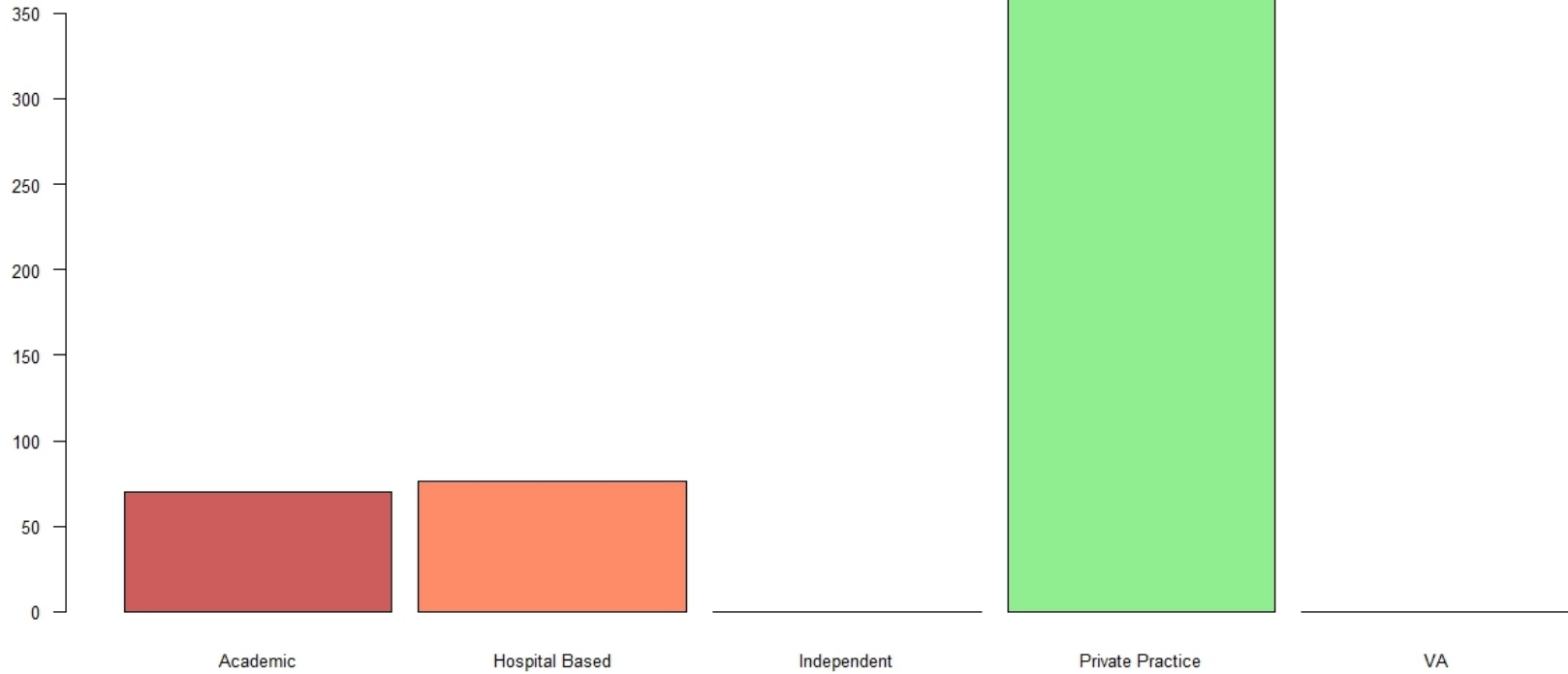
# Results

Site Type for ARO-CRO Organizations



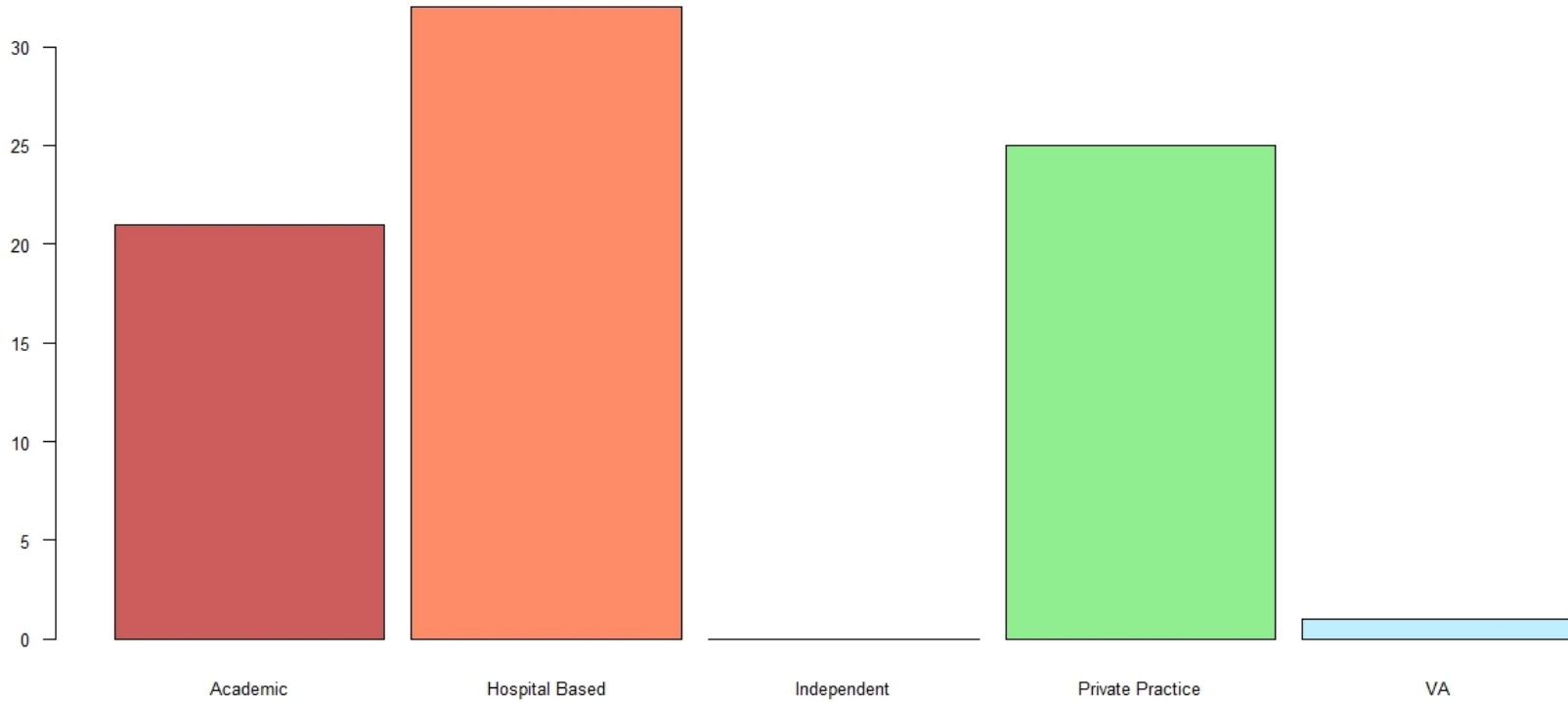
# Results

Site Type for Biotechnology Organizations



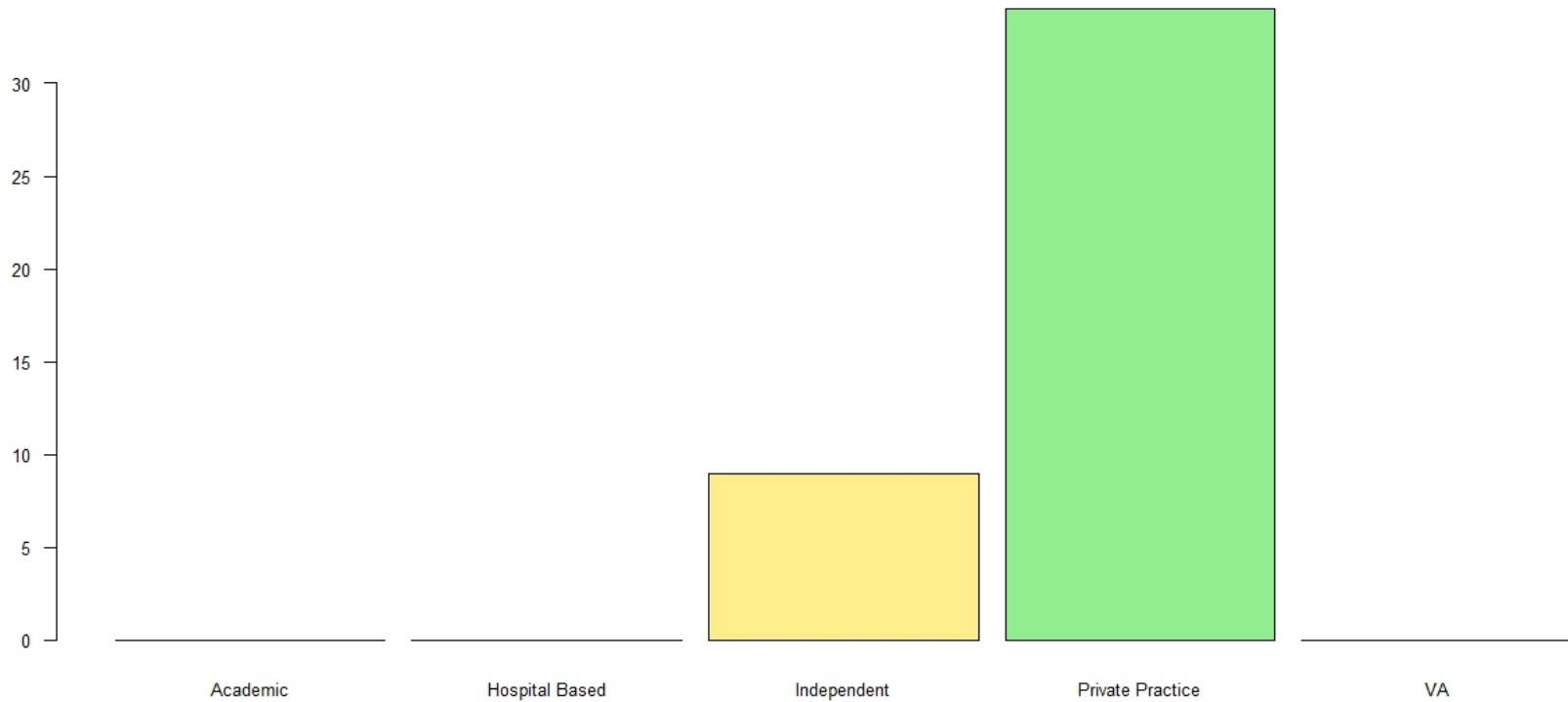
# Results

Site Type for Device Organizations



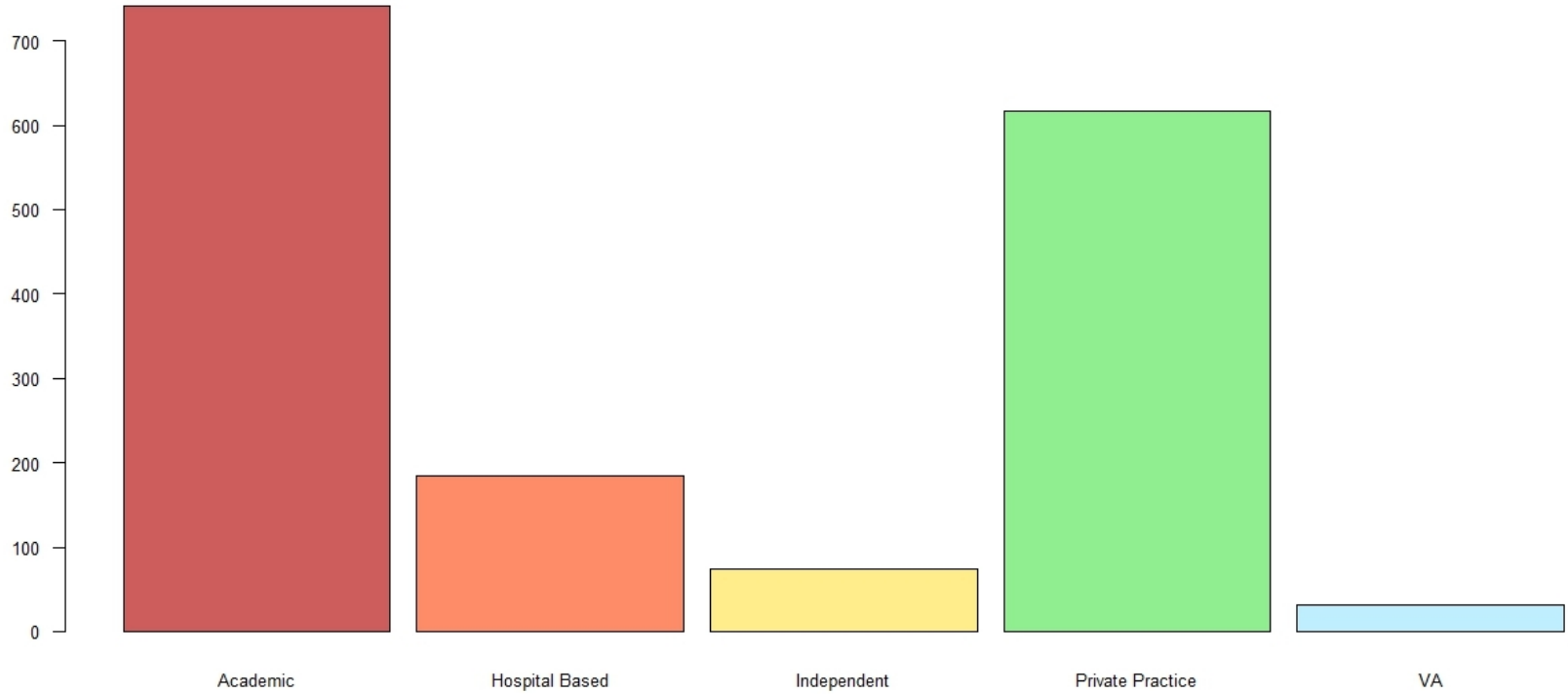
# Results

Site Type for Investigators Organizations



# Results

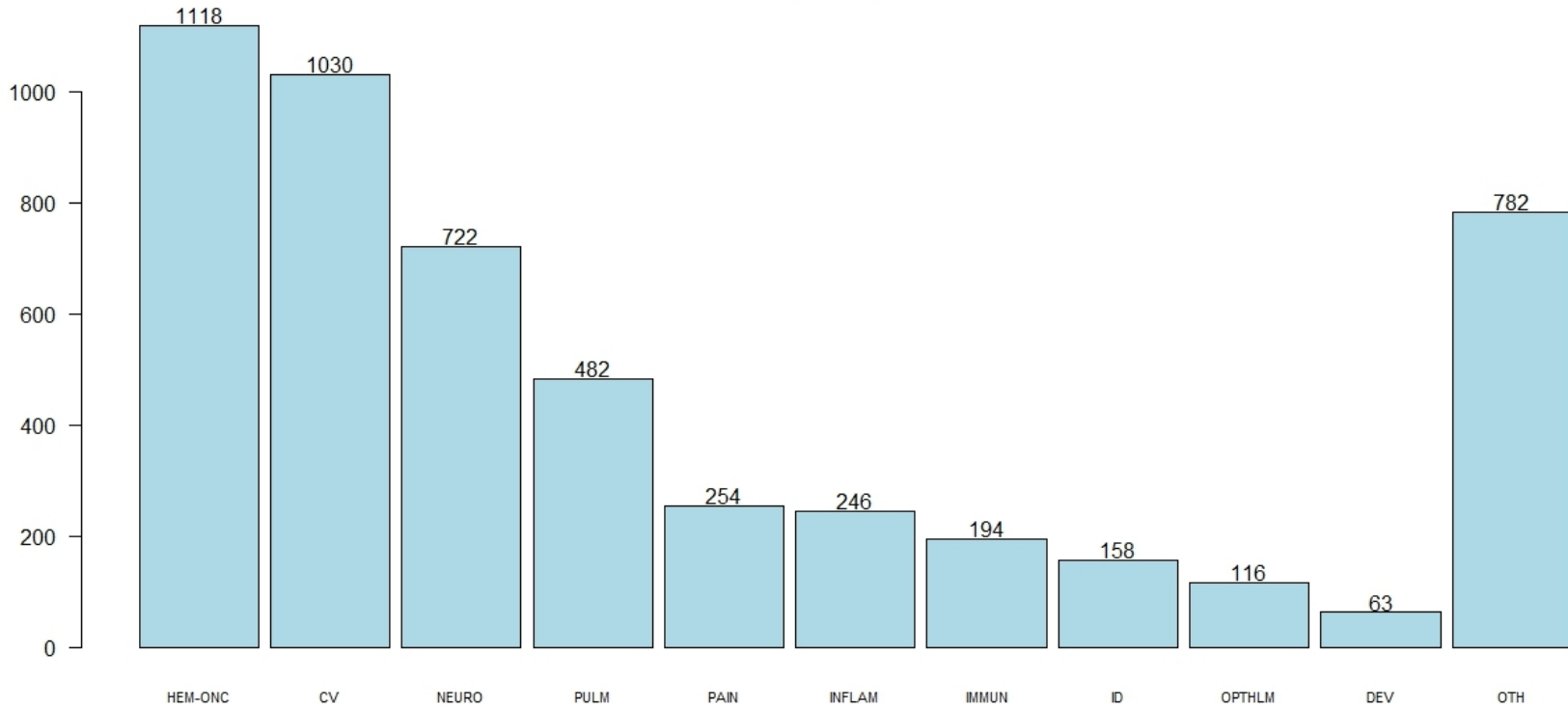
Site Type for Pharma Organizations



Which Therapeutic Areas are  
Presented in the Data?

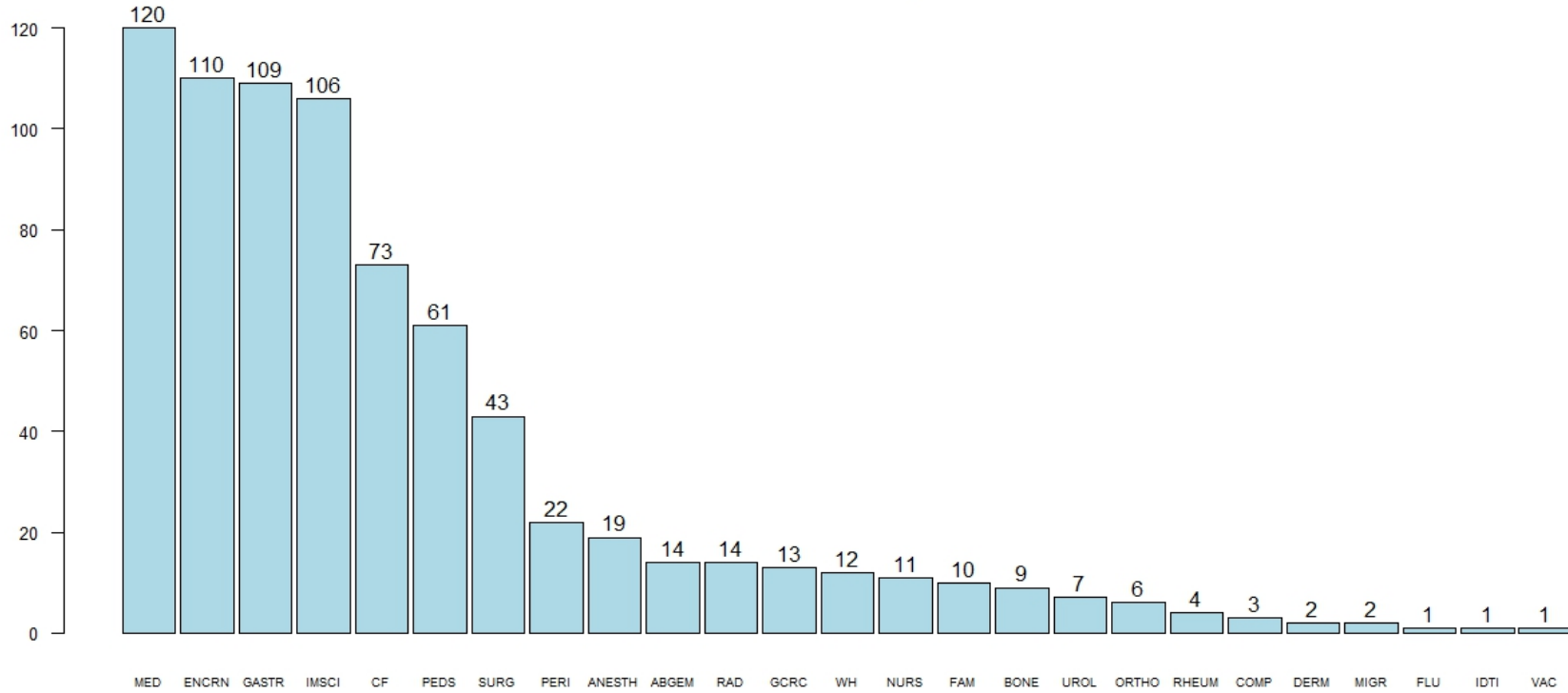
# Results

Totals by Therapeutic Area



# Results

Totals for Other Therapeutic Area



# Results

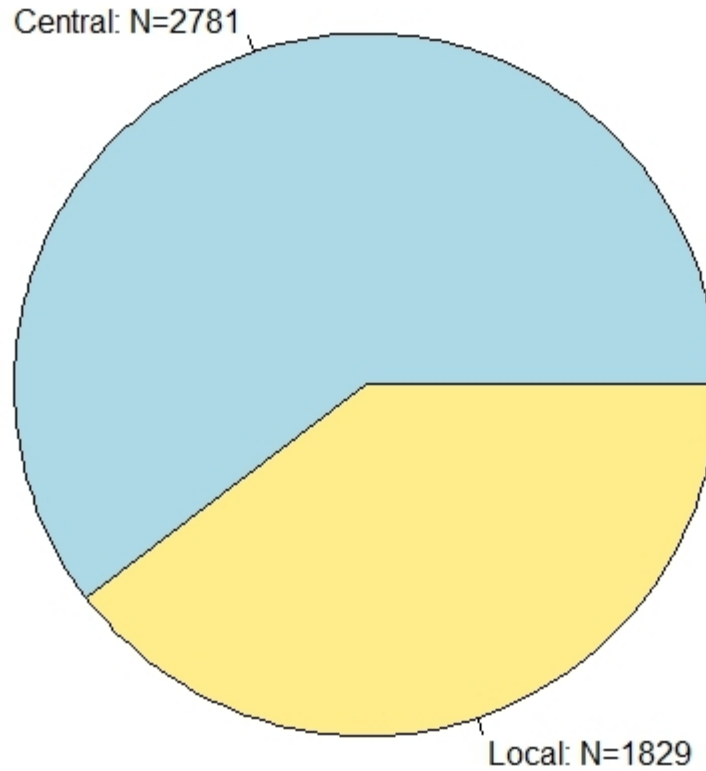
Other Therapeutic Area	Abbreviation	N
Medicine	MED	120
Endocrinology	ENCRN	110
Gastroenterology	GASTR	109
Immunoscience	IMSCI	106
Cystic Fibrosis	CF	73
Pediatrics	PEDS	61
Surgery	SURG	43
Perioperative	PERI	22
Anesthesiology	ANESTH	19
ABGEM	ABGEM	14
Radiology	RAD	14
GCRC	GCRC	13
Women's Health	WH	12
School of Nursing	NURS	11
Comm & Fam Med	FAM	10
Bone	BONE	9
Urology	UROL	7
Orthopedic Surgery	ORTHO	6
Rheumatology	RHEUM	4
Compassionate Use	COMP	3
Dermatology	DERM	2
Ped Migraine	MIGR	2
Flu	FLU	1
IDTI	IDTI	1
Vaccines	VAC	1



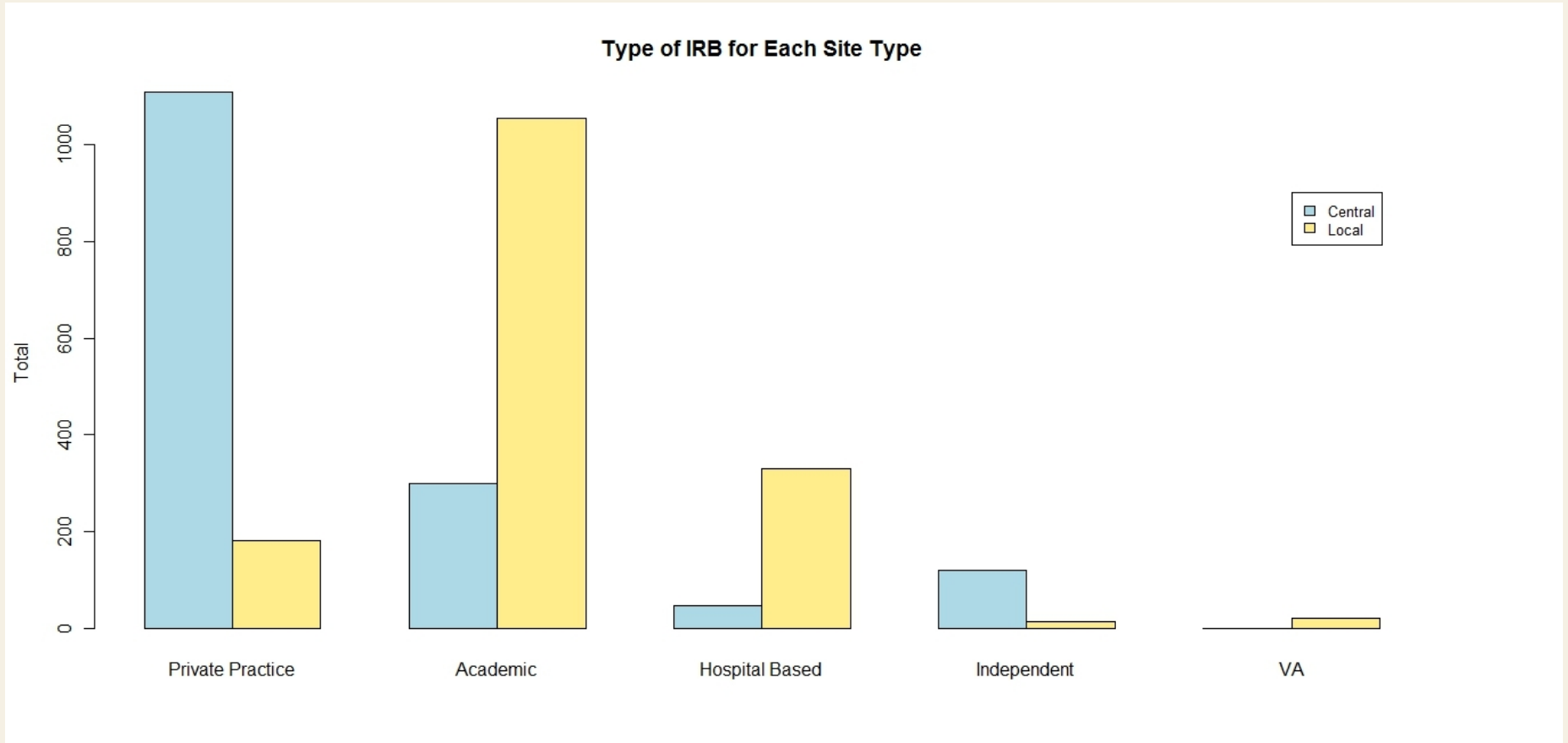
# What is the Distribution of Use of Central vs. Local IRBs?

# Results

All Data: IRB Type

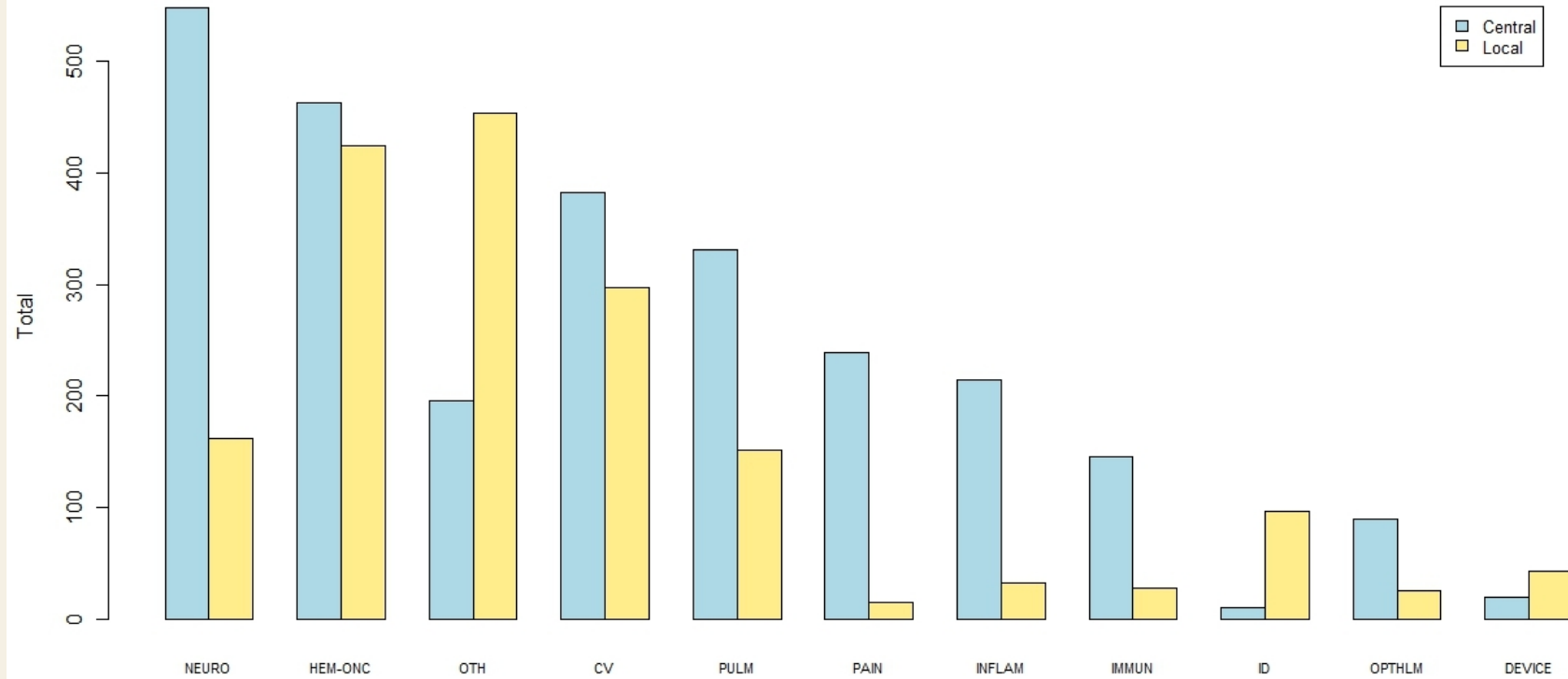


# Results



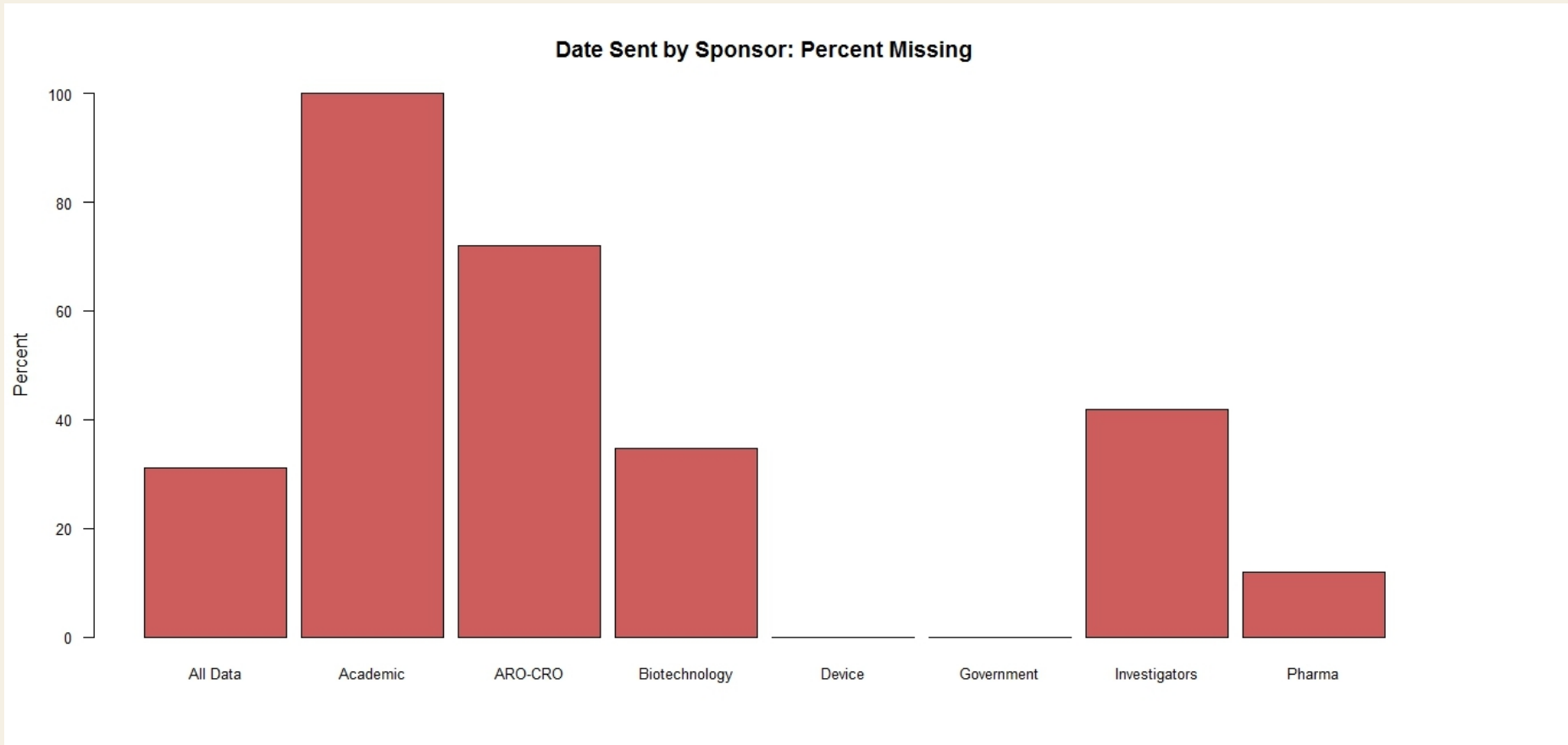
# Results

Type of IRB for Each Therapeutic Area



# What was the Extent of the Missing Data?

# Results

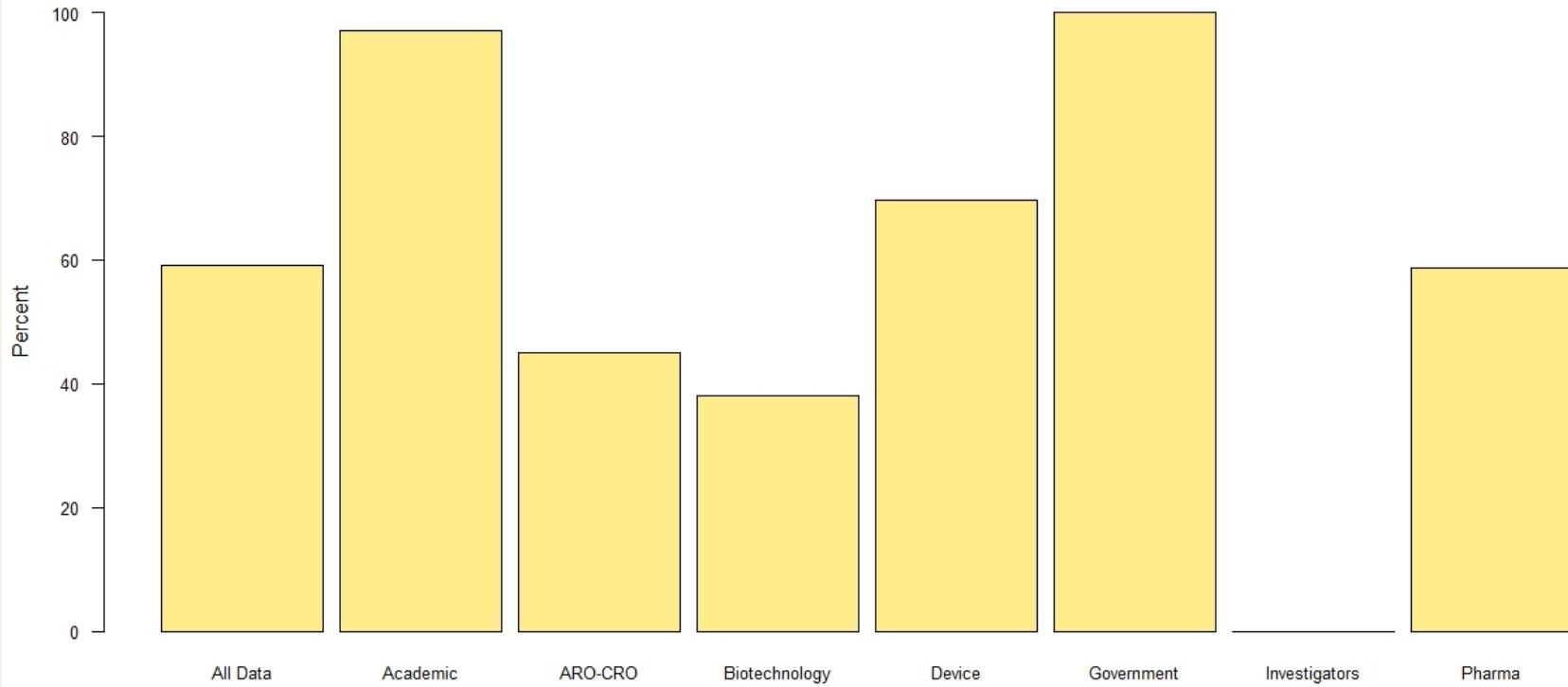


# Results



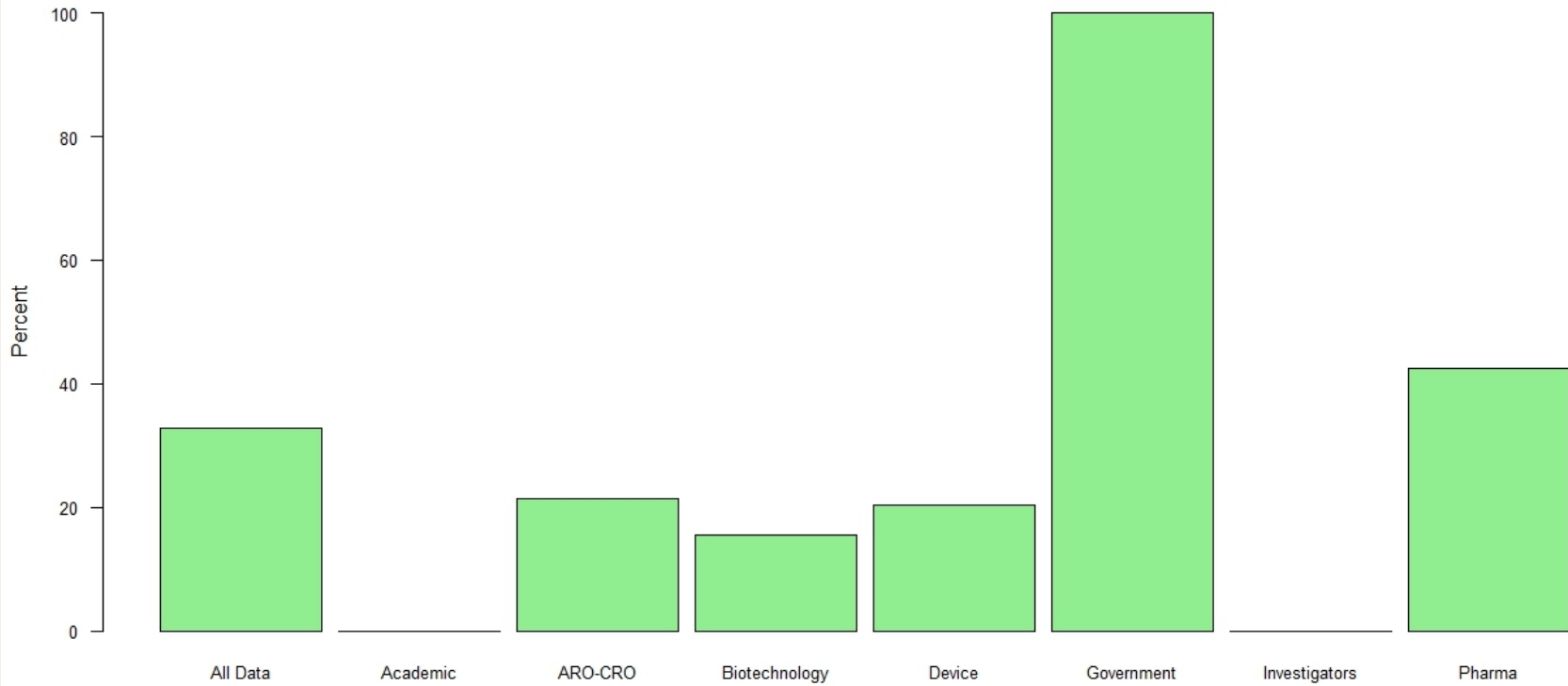
# Results

Date Submitted to IRB: Percent Missing

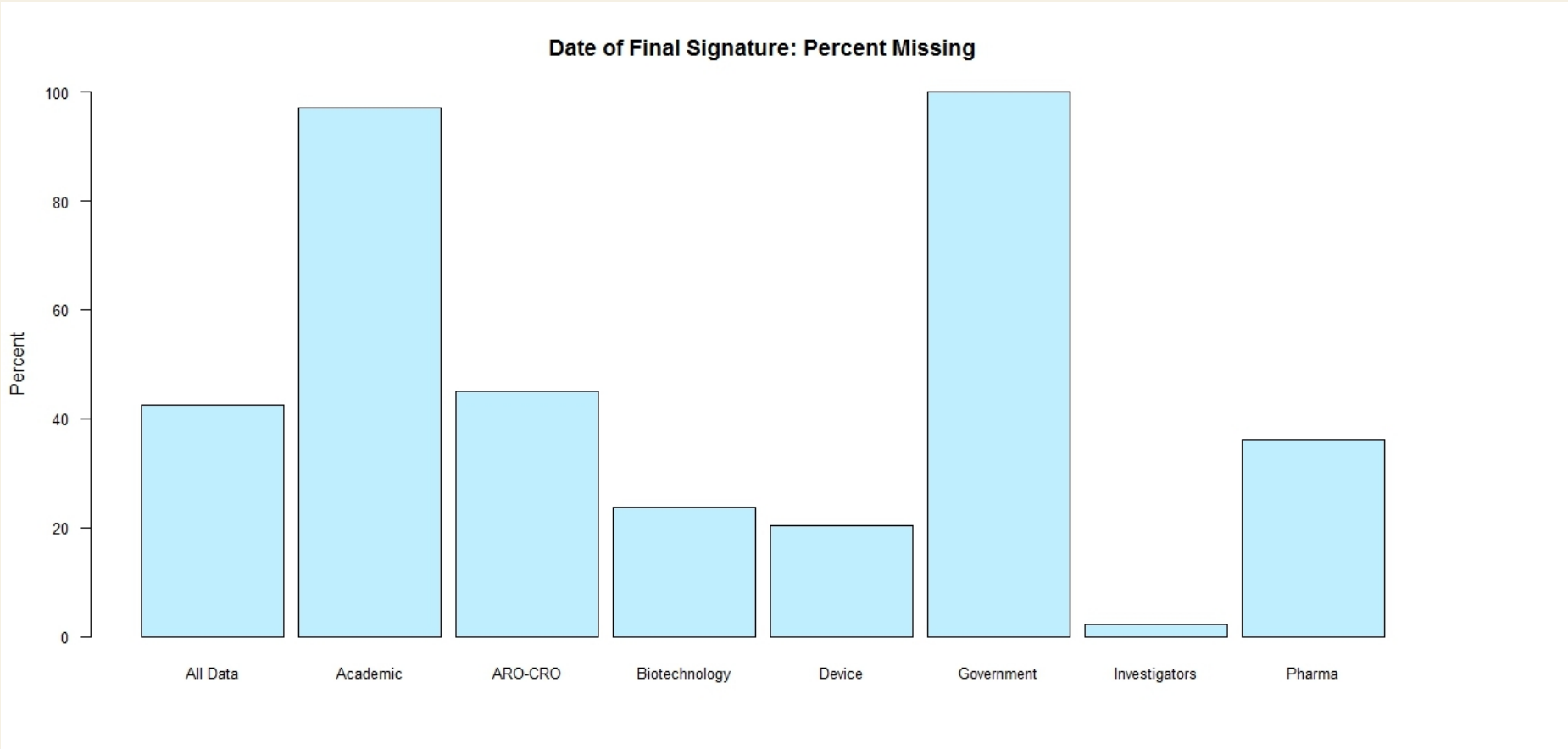


# Results

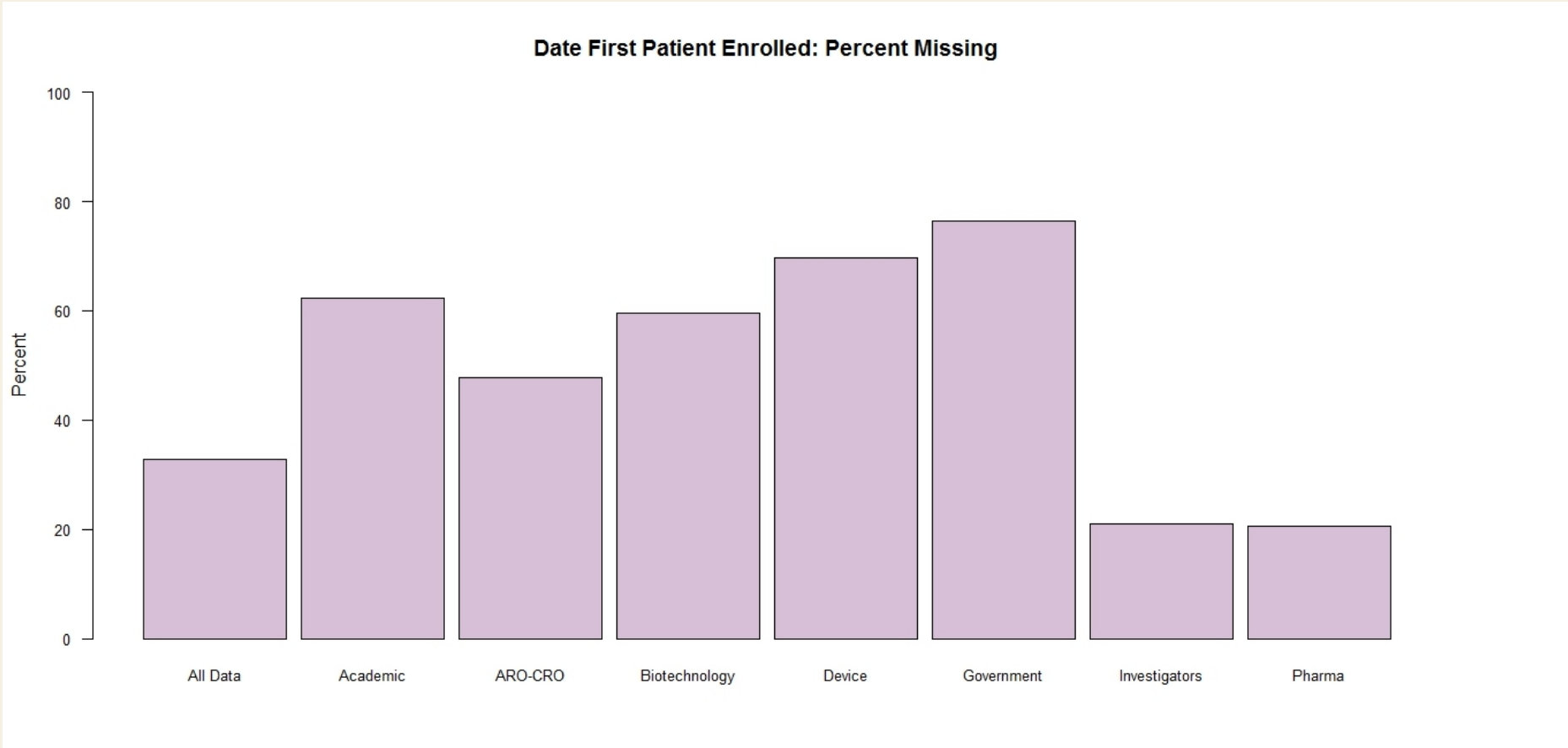
Date of IRB Final Decision: Percent Missing



# Results

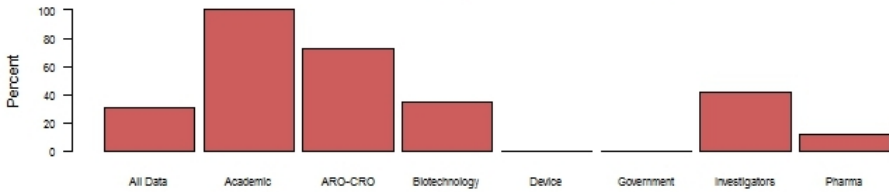


# Results

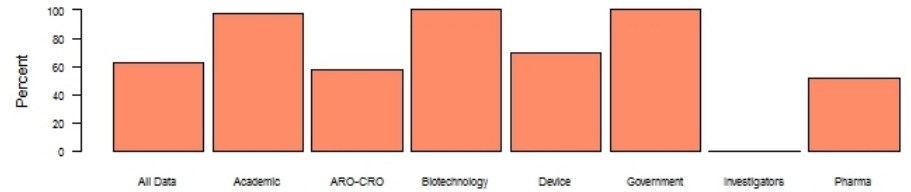


# Results

Date Sent by Sponsor: Percent Missing



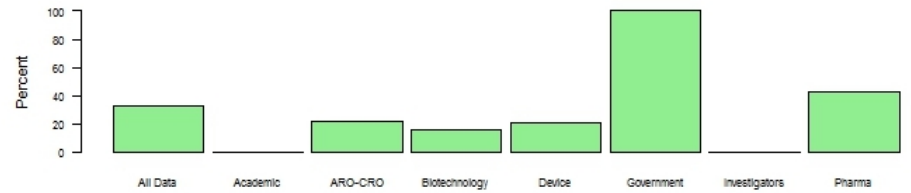
Date Received by Site: Percent Missing



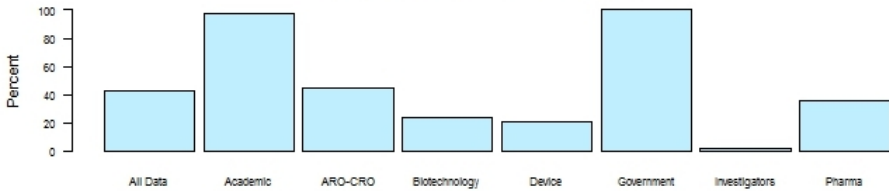
Date Submitted to IRB: Percent Missing



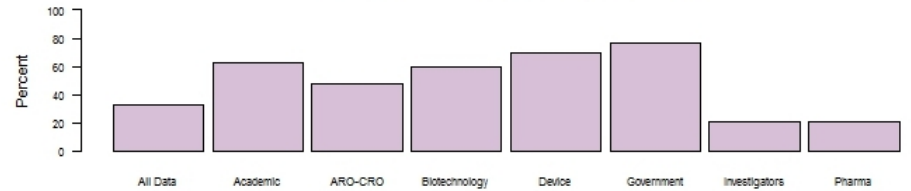
Date of IRB Final Decision: Percent Missing



Date of Final Signature: Percent Missing

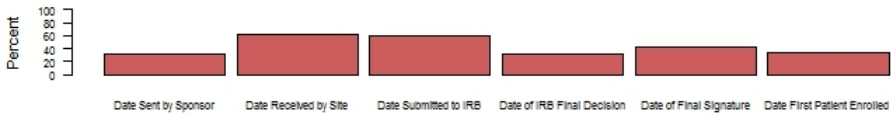


Date First Patient Enrolled: Percent Missing

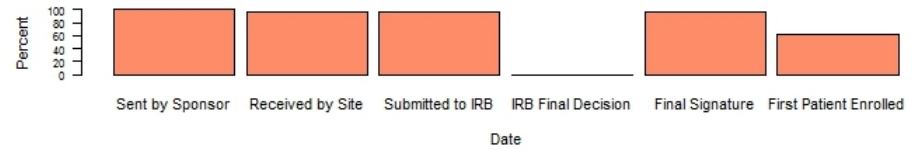


# Results

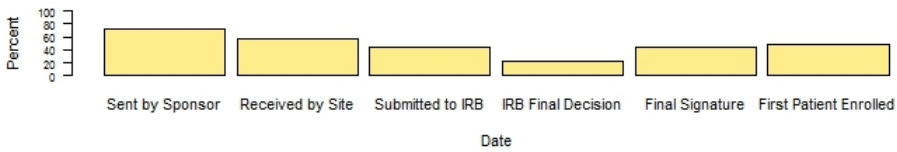
Overall Data: Percent Missing for Each Date Variable



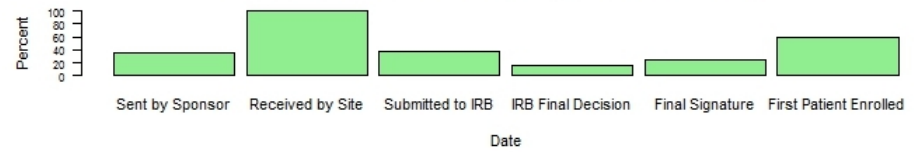
Academic Data: Percent Missing for Each Date Variable



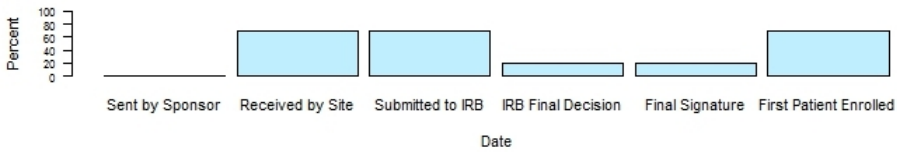
ARO CRO Data: Percent Missing for Each Date Variable



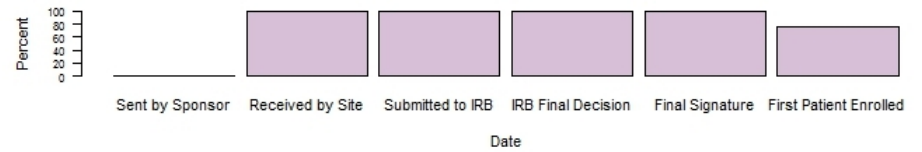
Biotechnology Data: Percent Missing for Each Date Variable



Device Data: Percent Missing for Each Date Variable



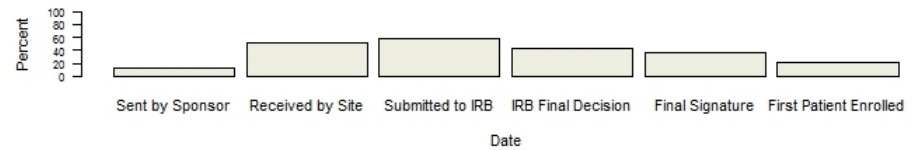
Government Data: Percent Missing for Each Date Variable



Investigators Data: Percent Missing for Each Date Variable



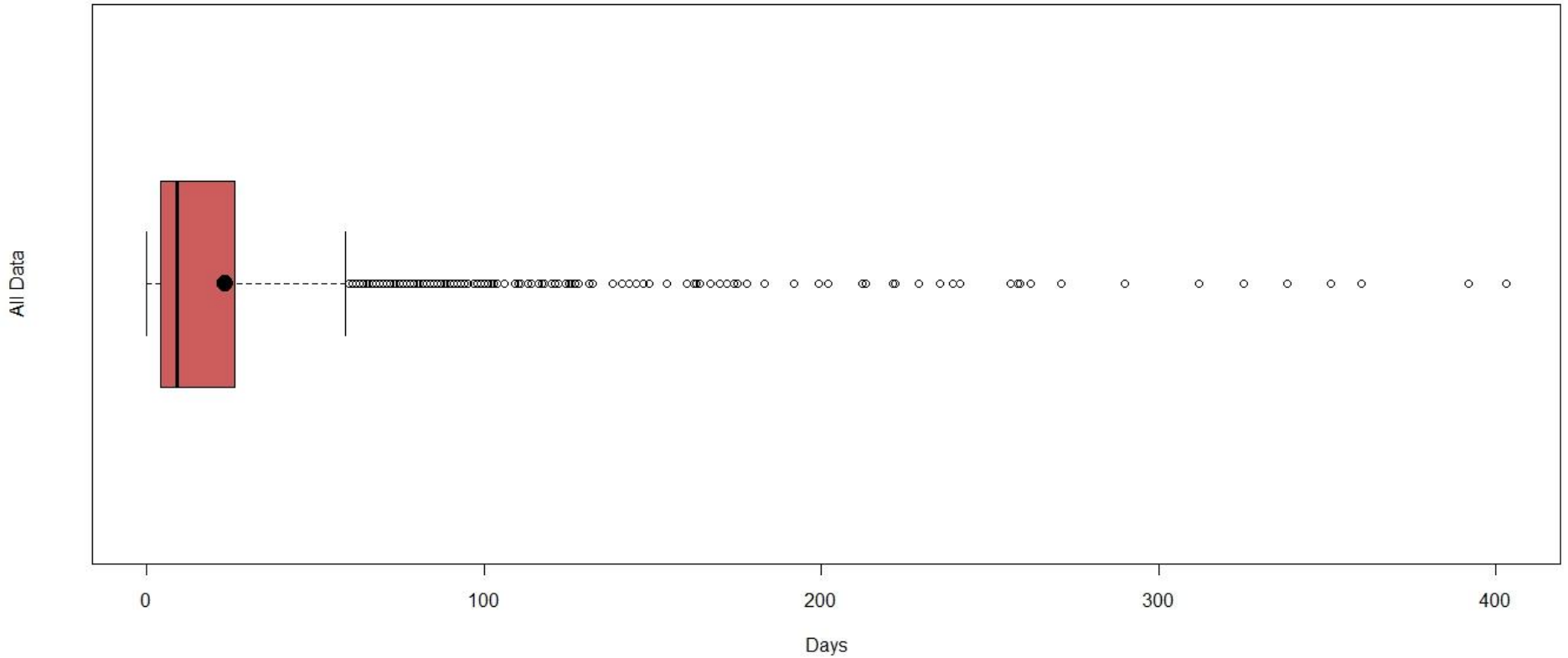
Pharm Data: Percent Missing for Each Date Variable



# What are the Cycle Time Results?

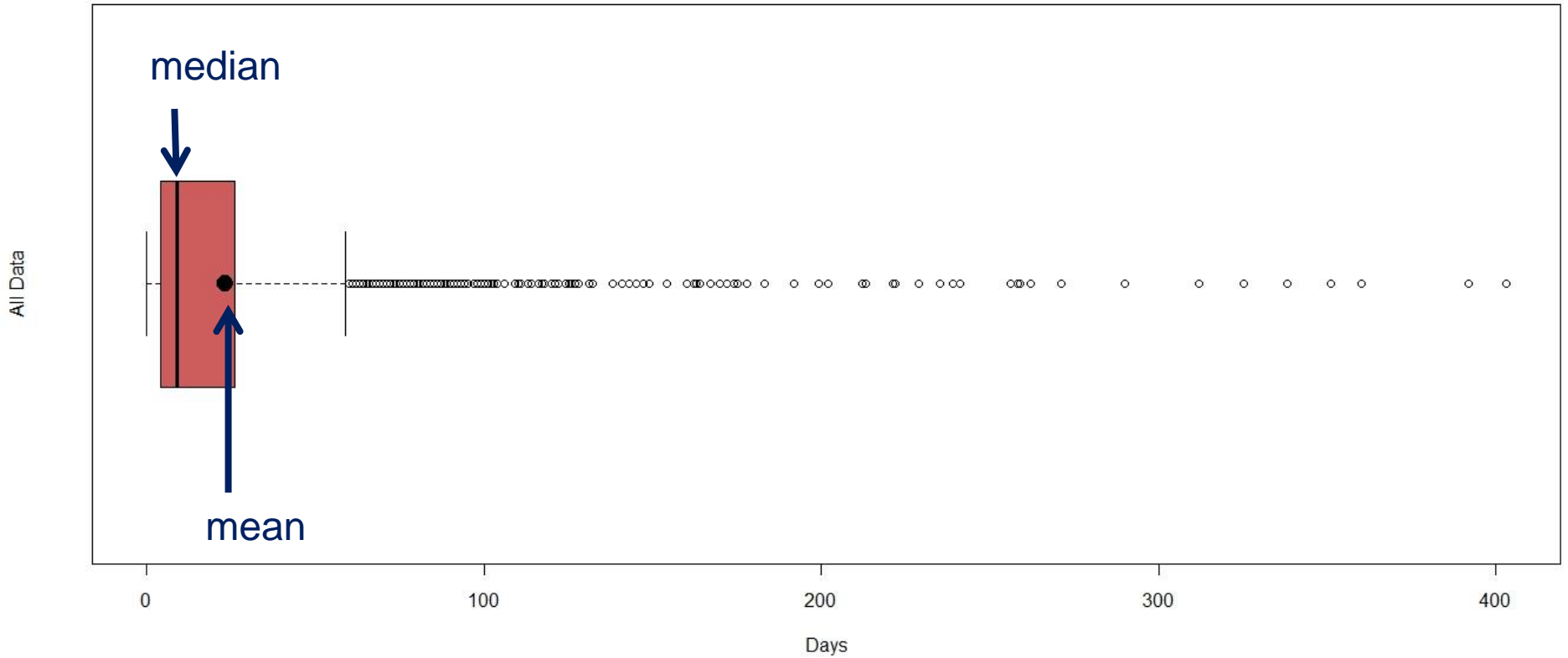
# Results

All Data: Date Protocol Submitted to IRB to Date of IRB Final Decision

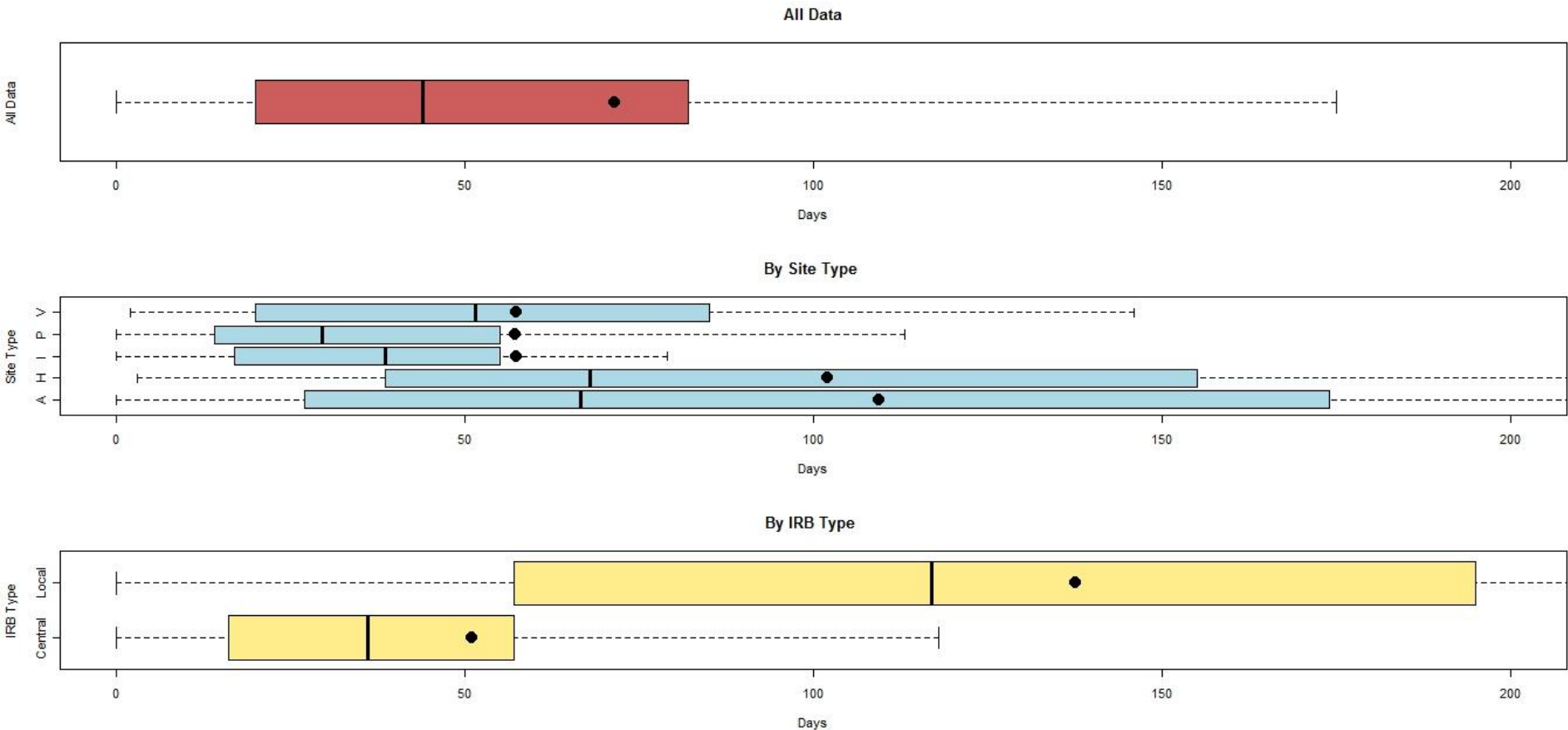


# Results

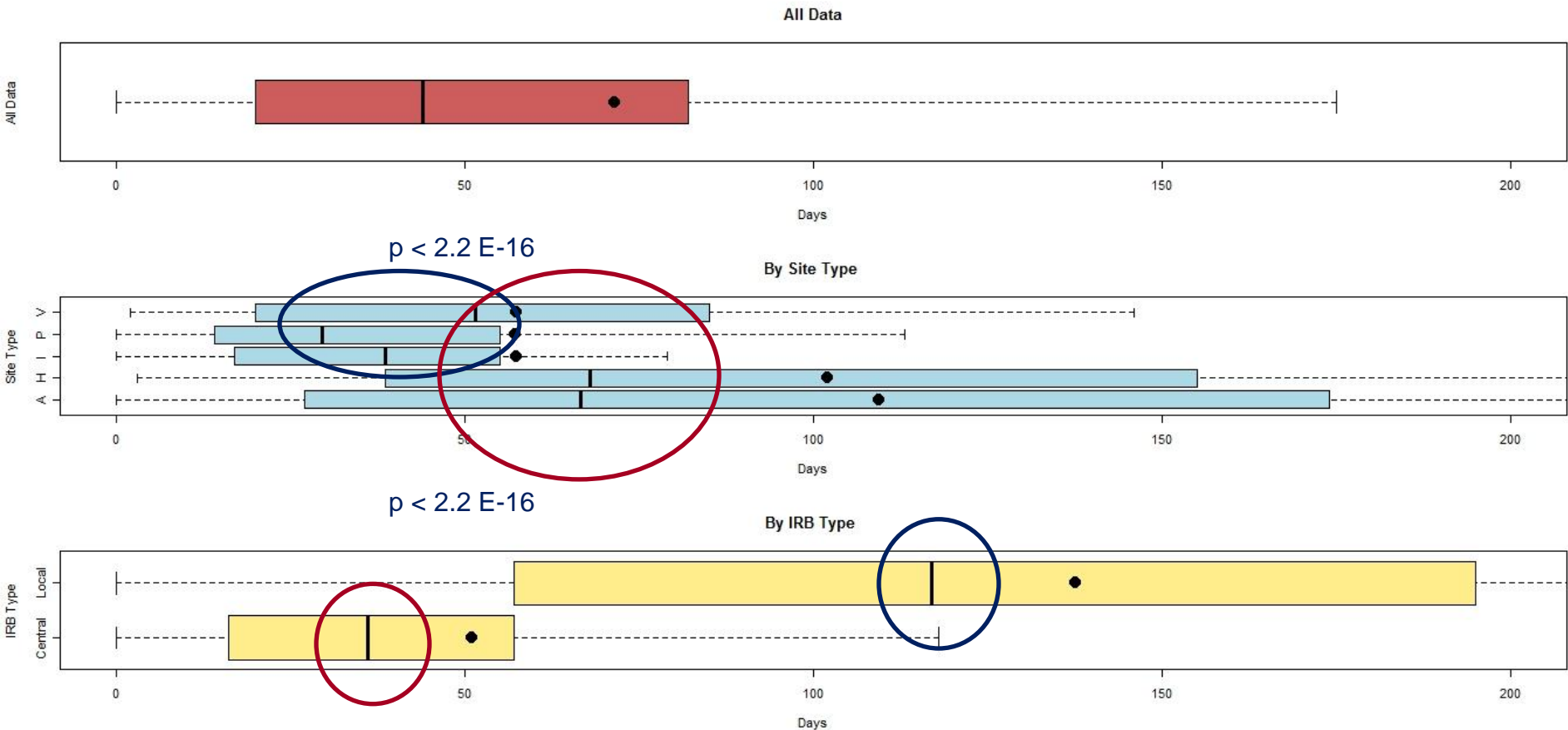
All Data: Date Protocol Submitted to IRB to Date of IRB Final Decision



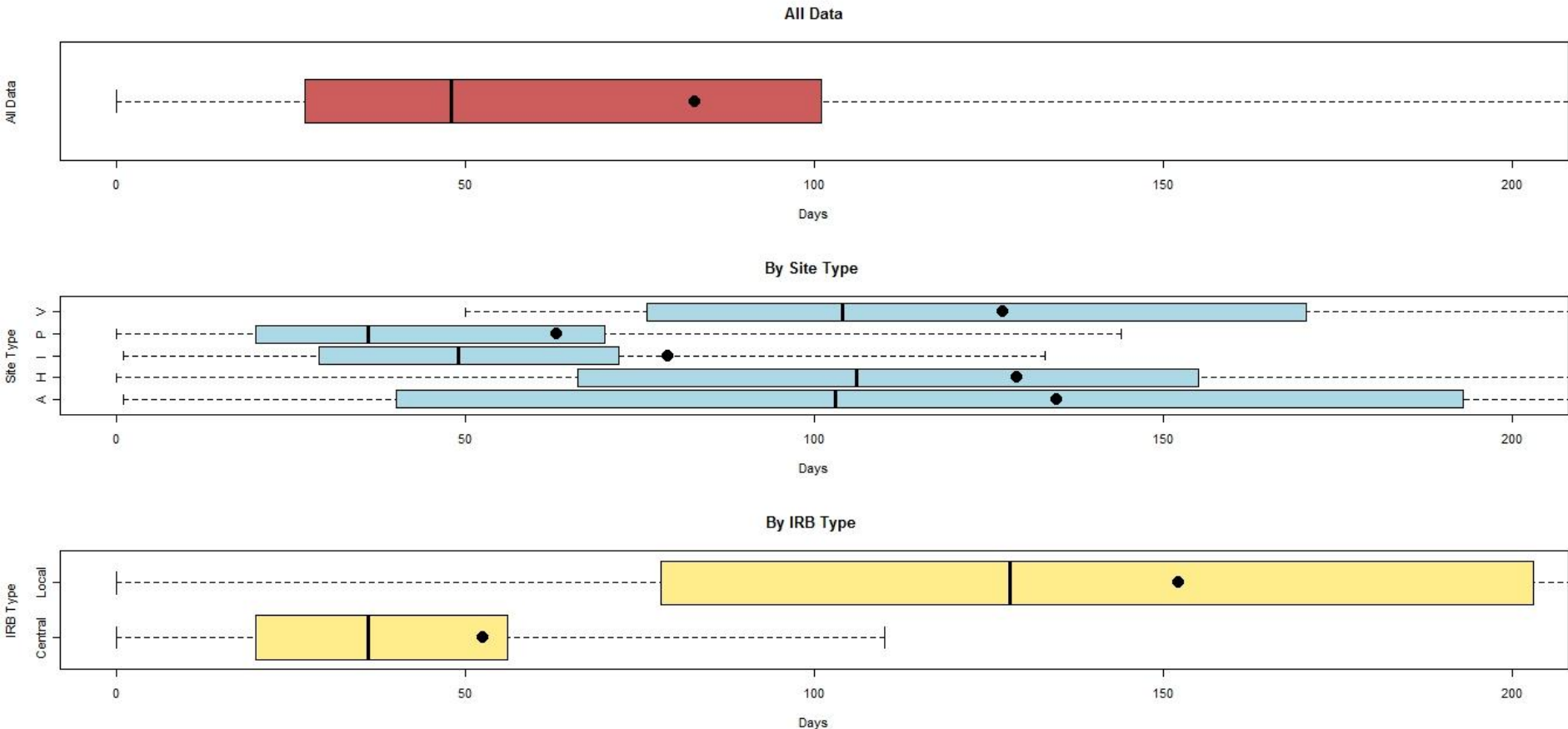
# Date Protocol Sent/Received to Date Protocol Submitted to IRB



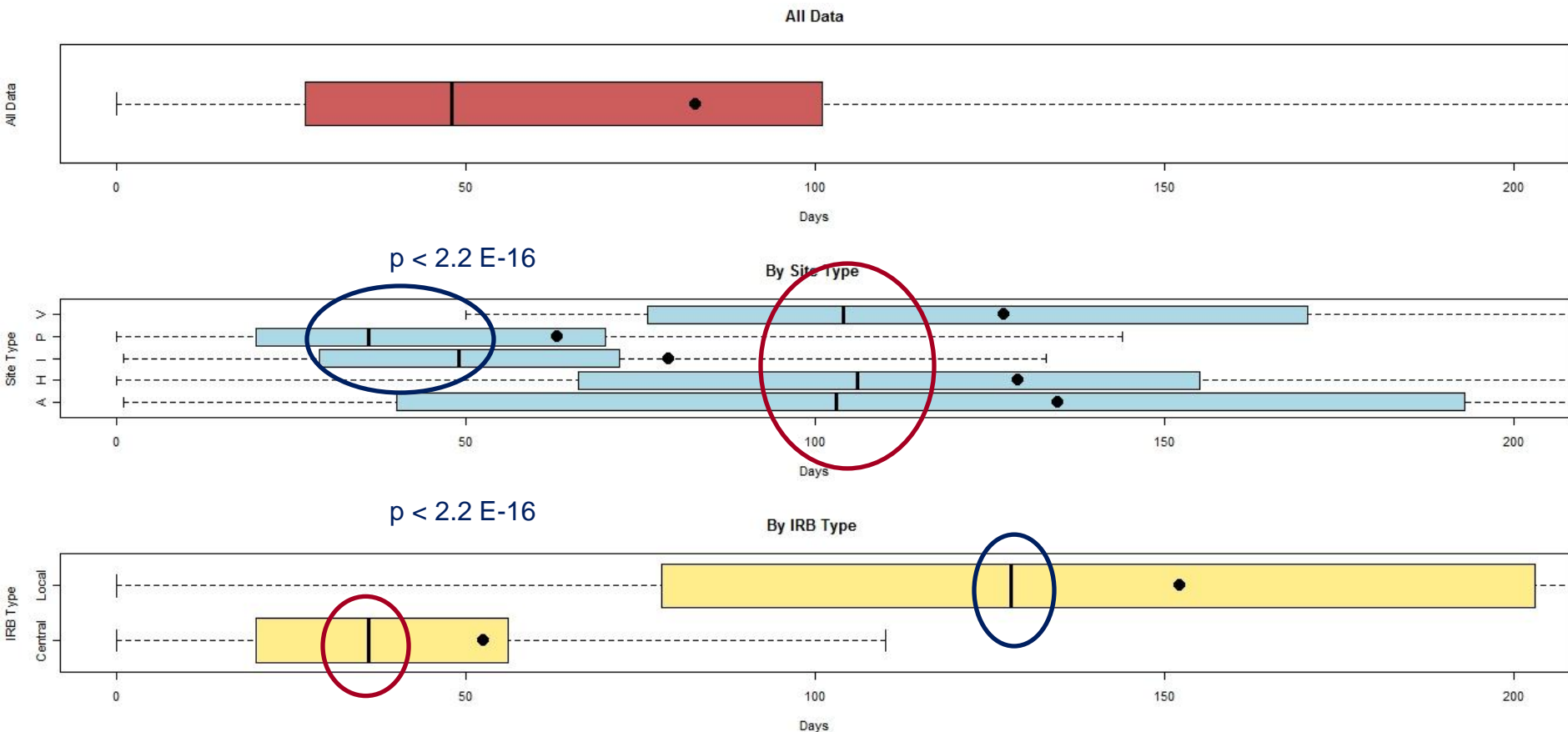
# Date Protocol Sent/Received to Date Protocol Submitted to IRB



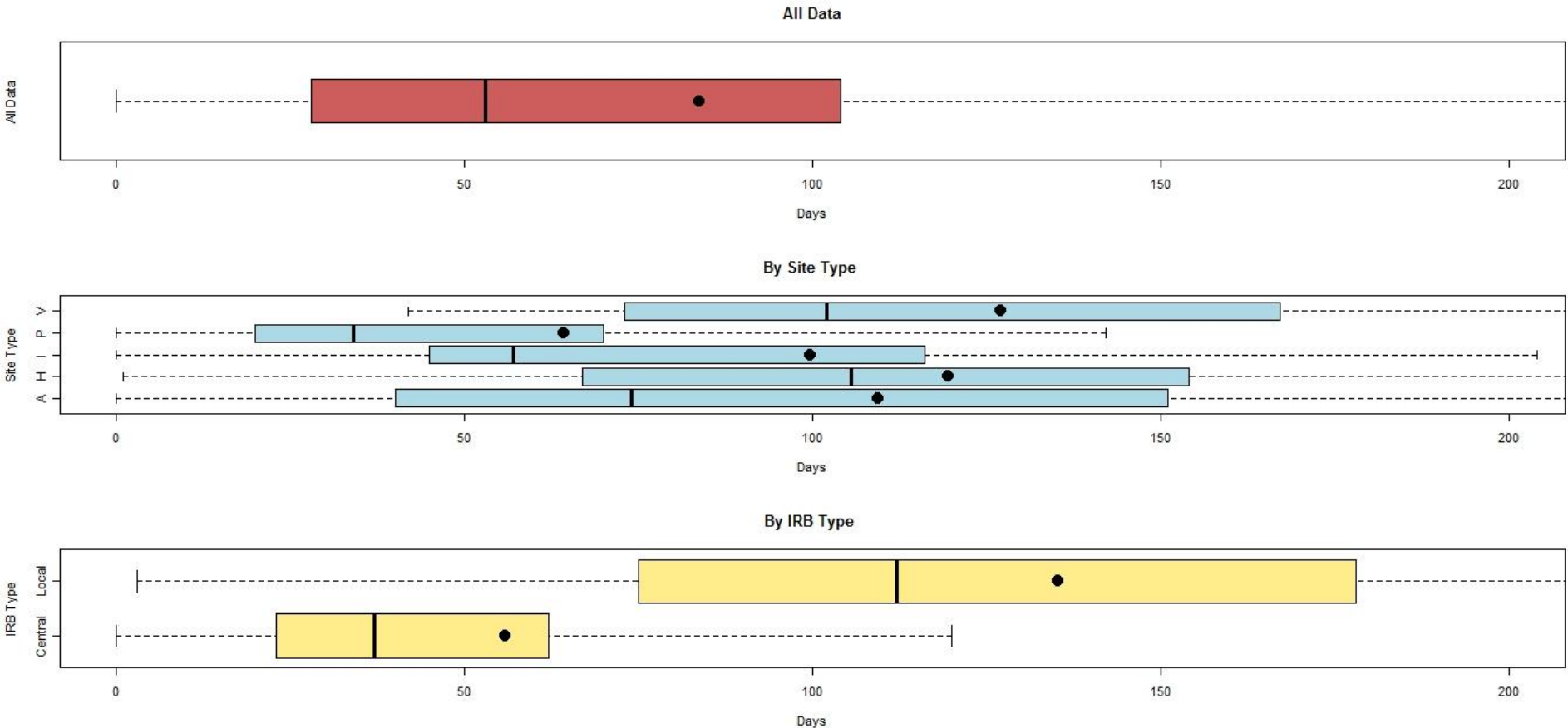
# Date Protocol Sent/Received to Date of IRB Final Decision



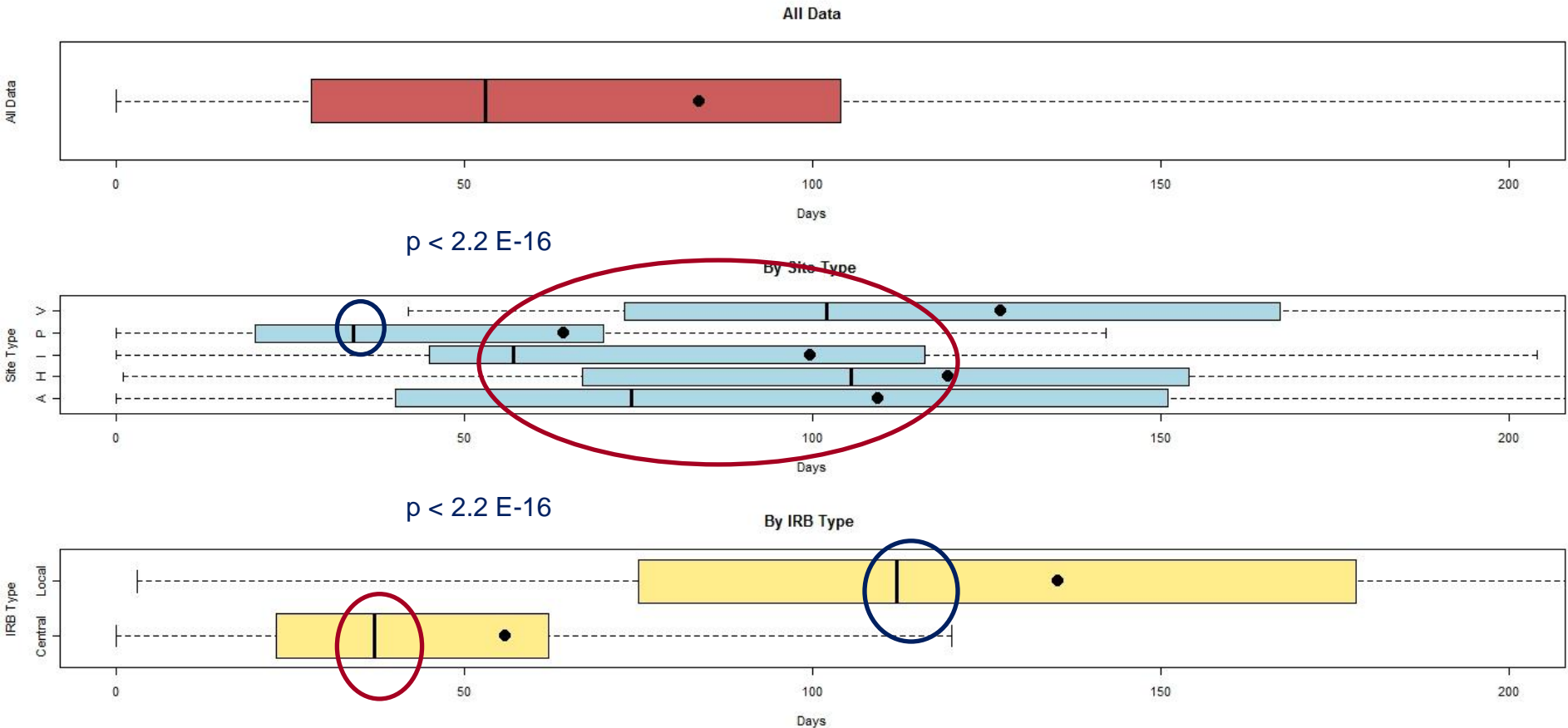
# Date Protocol Sent/Received to Date of IRB Final Decision



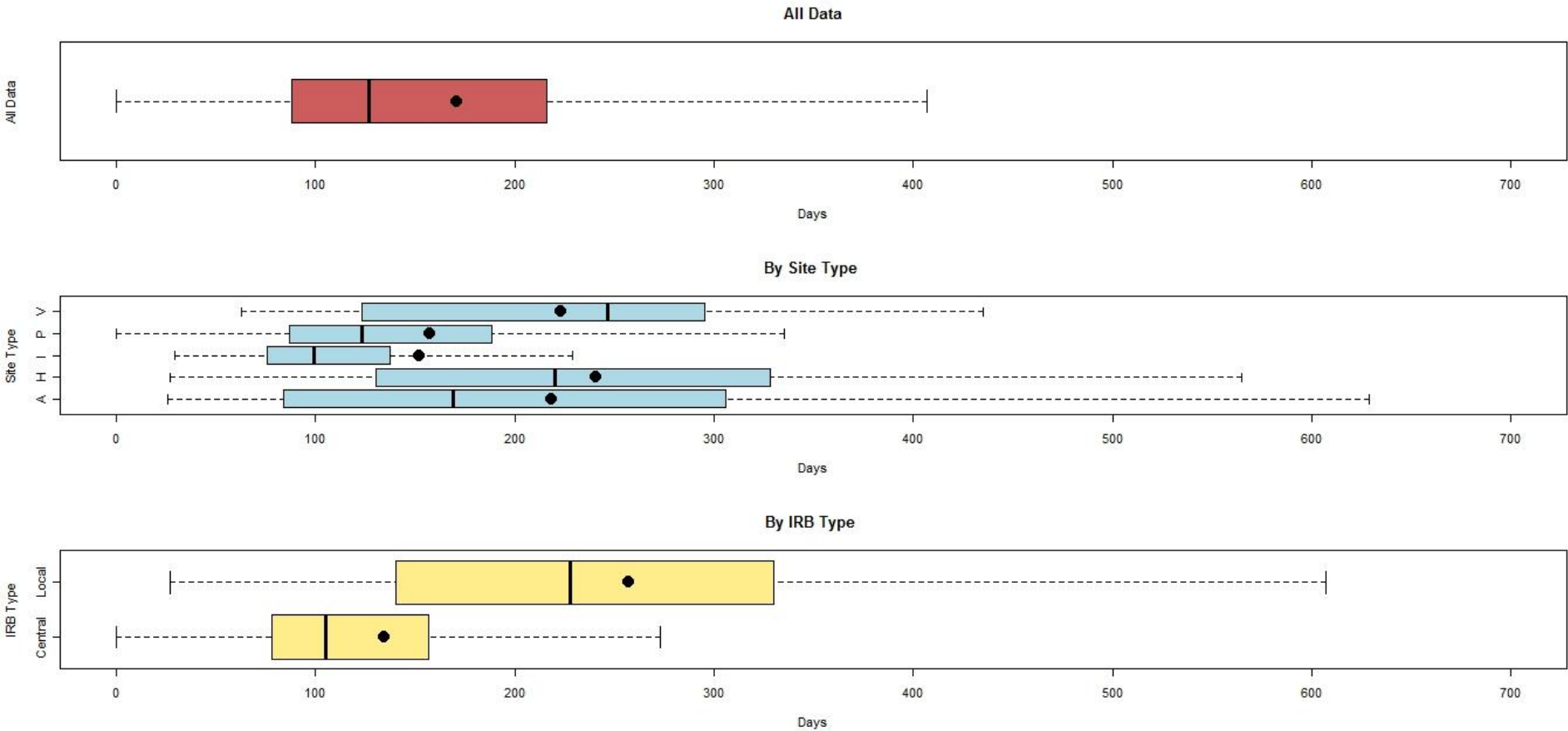
# Date Protocol Sent/Received to Date Contract Executed



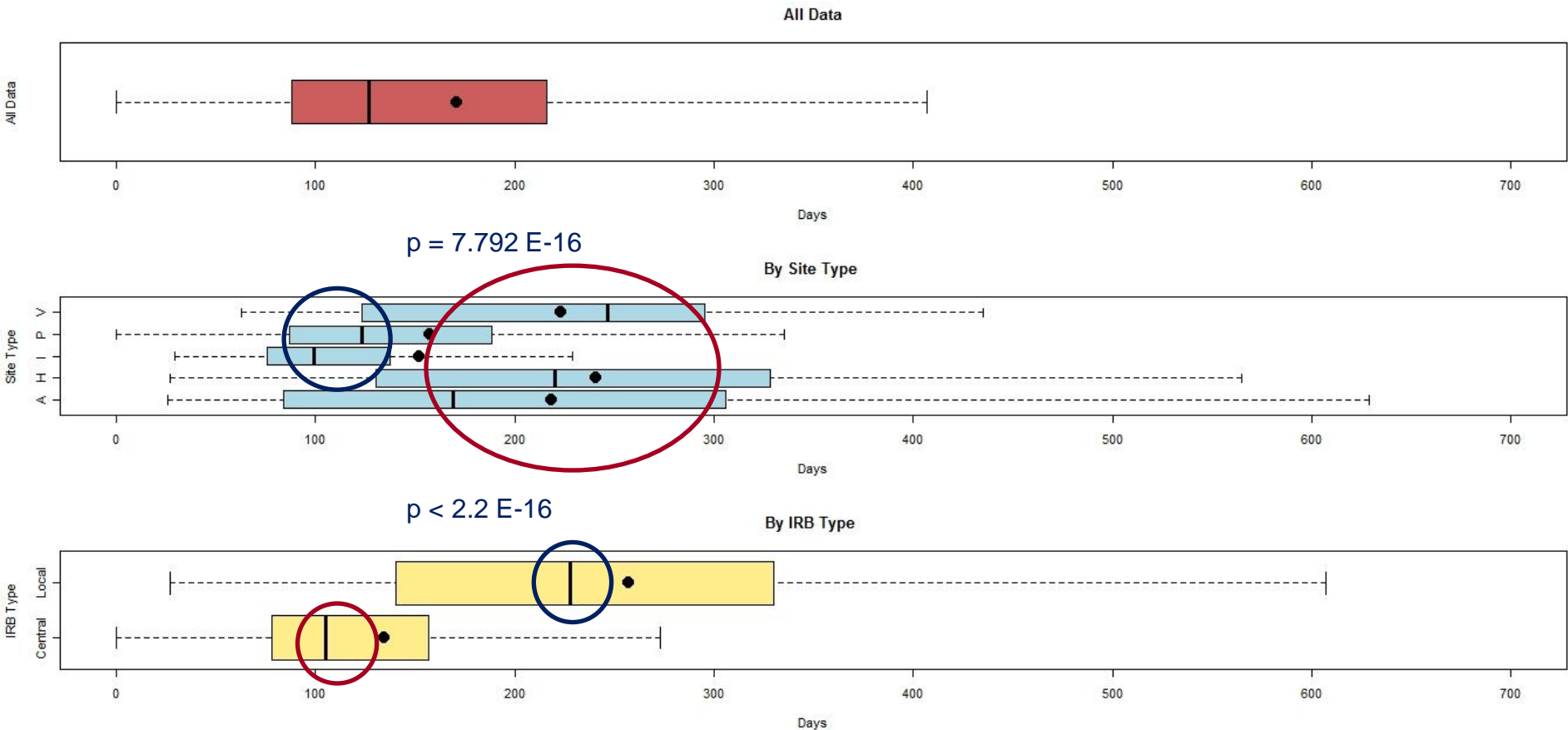
# Date Protocol Sent/Received to Date Contract Executed



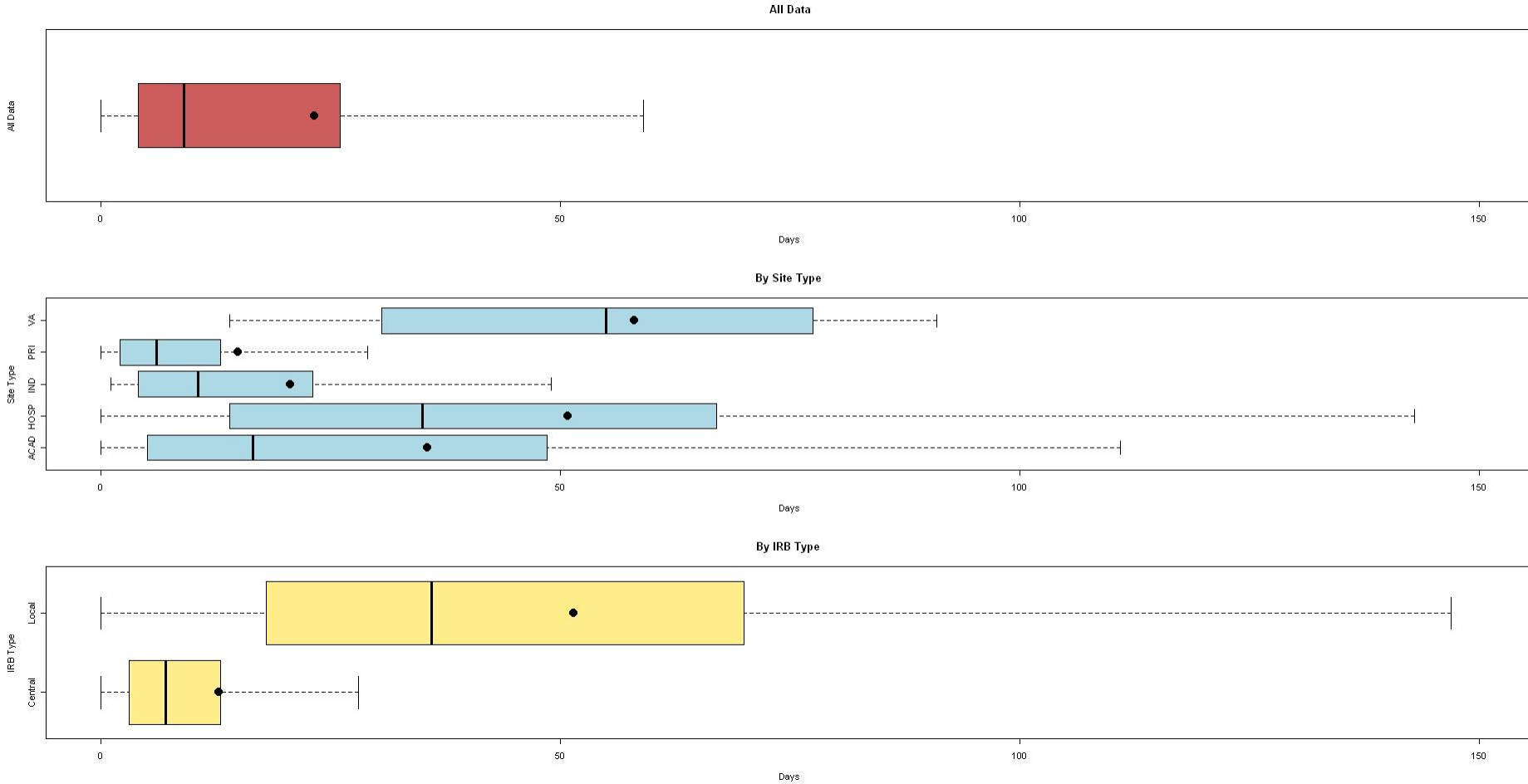
# Date Protocol Sent/Received to Date First Patient Enrolled



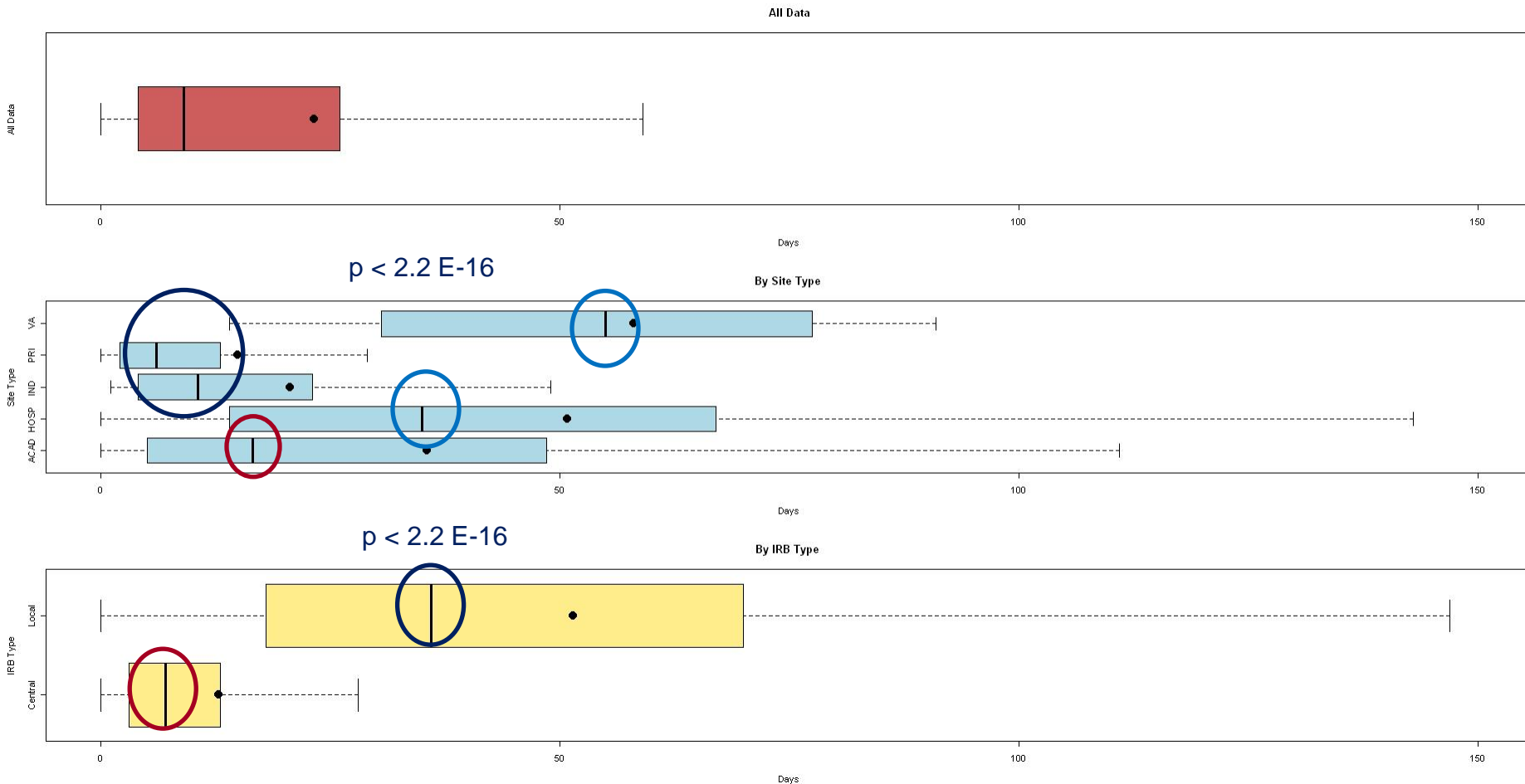
# Date Protocol Sent/Received to Date First Patient Enrolled



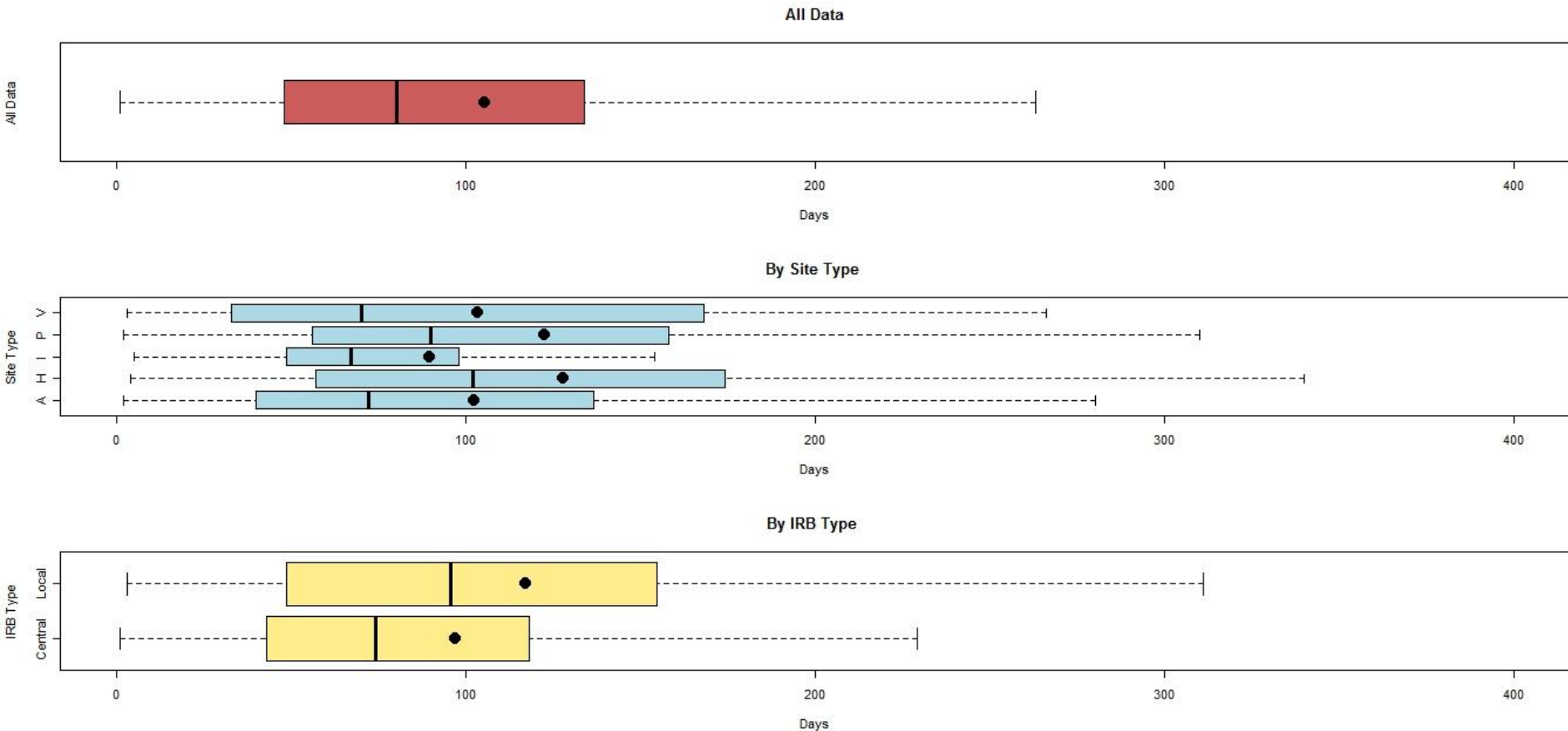
# Date Protocol Submitted to IRB to Date of IRB Final Decision



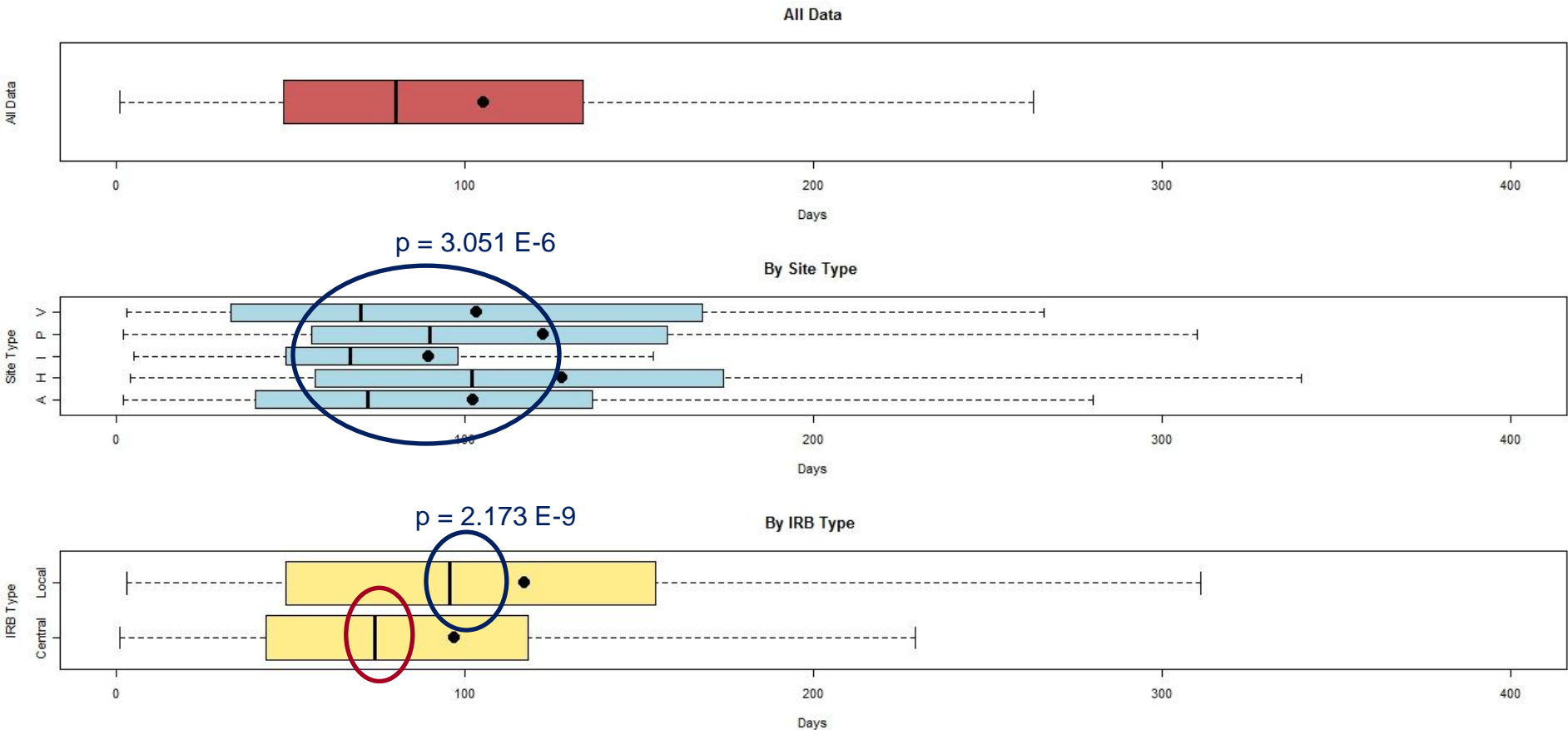
# Date Protocol Submitted to IRB to Date of IRB Final Decision



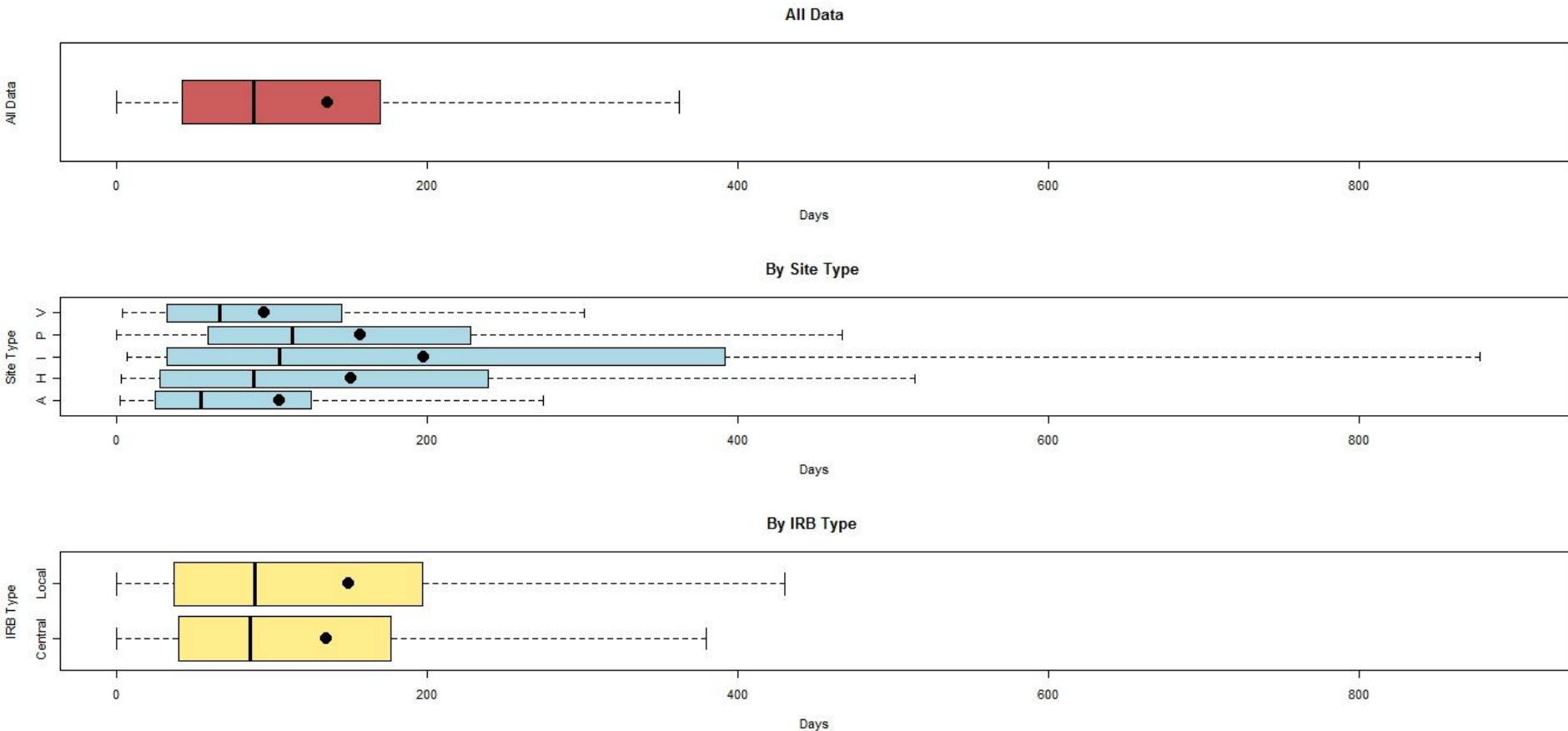
# Date of IRB Final Decision to Date First Patient Enrolled



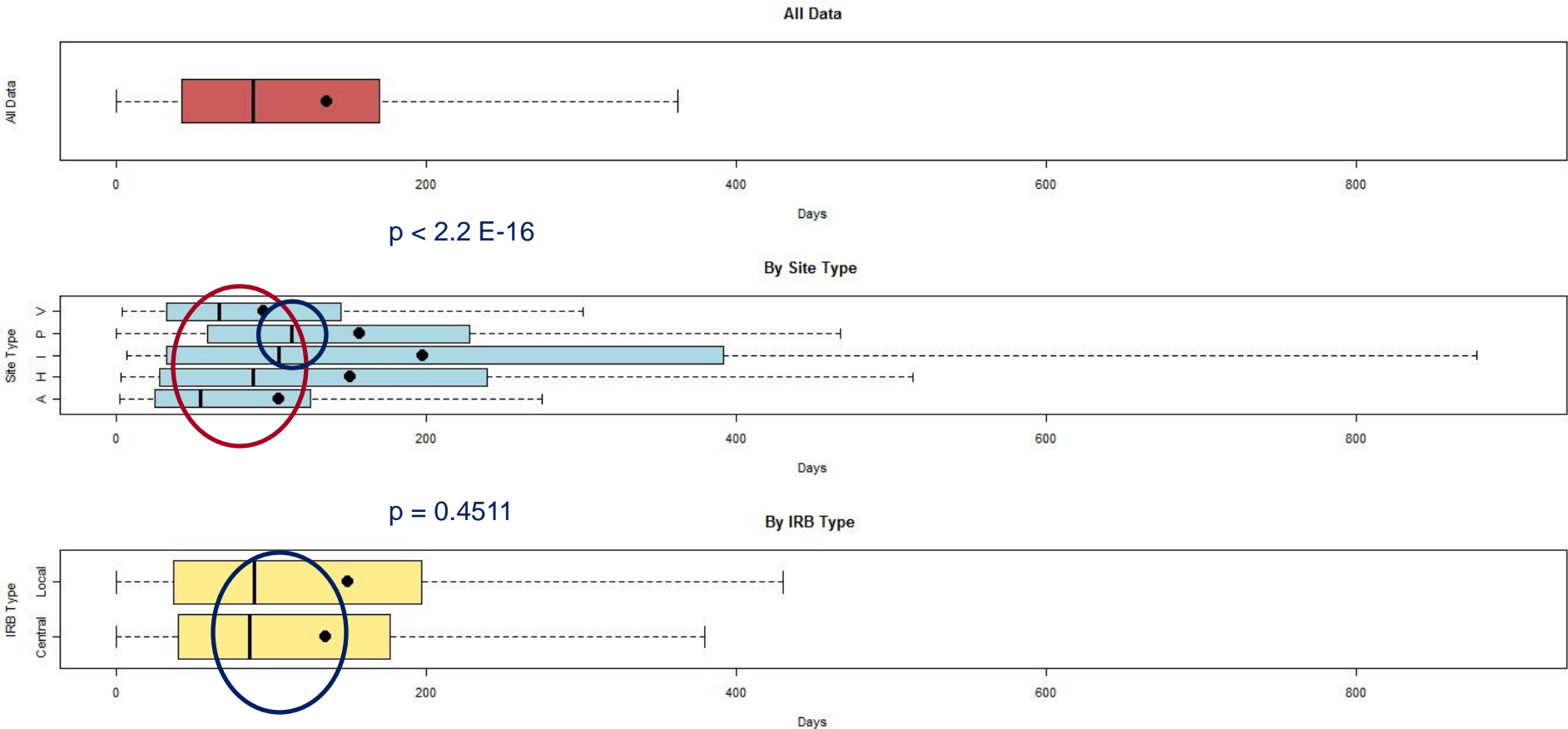
# Date of IRB Final Decision to Date First Patient Enrolled



# Date Contract Executed to Date First Patient Enrolled



# Date Contract Executed to Date First Patient Enrolled



# Limitations

- Sample size small in many areas
- Limited data submitted by sites themselves
- Lack of standard data elements definitions
  - ◆ Data elements definitions provided were not granular enough
- Goal of data analysis - to show variation and themes
- Large percentage of missing data
  - ◆ Data element(s) not routinely captured by submitting organization
  - ◆ Derived data where possible

# Lessons Learned

- Need to develop understanding of a common workflow
- Need to develop enterprise wide standard definitions for data elements
- Need to agree on which cycle time metrics are critical for monitoring study start-up
- Avoid distraction of site-specific processes
- Must provide the “what’s in it for me” explanation to facilitate data collection
- Must be as granular as possible in providing definitions
- Need to provide drop down option lists when possible
- Need more non-date data collected to facilitate identification of processes that could be responsible for cycle time variation

# Questions?