

EVALUATION OF EFFICIENCY IN THE ACTIVATION AND ACCRUAL OF
INTERVENTIONAL CLINICAL TRIALS AT CANCER CENTERS

by

Wendy Rose Tate

Copyright © Wendy Tate 2016

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF PHARMACEUTICAL SCIENCES

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2016

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Wendy Tate, titled "Evaluation of Efficiency in the Activation and Accrual of Interventional Clinical Trials at Cancer Centers" and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

_____ Date: March 2, 2016
Ivo Abraham

_____ Date: March 2, 2016
Marion Slack

_____ Date: March 2, 2016
Terri Warholak

_____ Date: March 2, 2016
Robin Harris

_____ Date: March 2, 2016
Lee Cranmer

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

_____ Date: March 2, 2016
Dissertation Director: Ivo Abraham

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of the requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that an accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: Wendy Rose Tate

ACKNOWLEDGEMENTS

I've been saying for years that, "It takes a village to put Wendy through grad school." It truly did, and now that I am completing this journey, I want to thank my village. Thank you to my professors and committee, Drs. Ivo Abraham, Marion Slack, Terri Warholak, Robin Harris, and Lee Cranmer. Thank you for your mentorship, wisdom, and teaching over the last five years.

I have been blessed to have several mentors and professionals during my career who have invested in me as an employee, a student, and a person. Special thanks go to Dr. Sheryl Wurl, who has always provided sage advice, empathy, as well as telling me when to buck up and conquer. Thank you Dr. Grant Skrepnek, who was willing to have a prospective student drop by his office to ask a few questions and spent two and a half hours talking about clinical research, regulations, and changing the FDA approval process, inspiring me to join the program. Thank you to Dr. Anne Cress, Dr. David Alberts, and Ms. Samantha Skrepnek for believing in my ability to assist in running a cancer clinical trials office. Many thanks go to Dr. Cranmer for giving me the opportunity to learn the issues of clinical trials operations firsthand and allowing me to be his lieutenant and research partner. We have done battle, and this project is a direct result of the hours of brainstorming on how to fix the issues surrounding clinical trial administration and bettering human health through quality and timely research.

I am indebted to the sixteen institutions who provided data for this project. The clinical trial office administrators had no requirement to take the time and resources to fulfill my requests, yet they did and they answered my (sometimes numerous) follow-up questions enthusiastically. Special thanks go to Brenda Pierson, who provided valuable editing input on this document.

Thank you to my amazing family and friends who have provided love, support, meals, drinks, babysitting services, shoulders to cry on, impromptu parties, inspirational talks, prayers, needed (and sometimes forced) camping trips, card games, and their complete faith that I could complete this. Thank you Mom and Dad, Mark and Lynnette, Brenda and Brian, Grandpa and Grandma (miss you). Thank you Mariette, Amber, Josh (JJ), Ed, and Devonnie. Thank you Jill for being my partner in crime through this program; I truly wouldn't have survived coursework and comprehensive exams without you.

Saving the best for last, I do not have enough thanks for my crew: Michael, Allyson, Caleb, and Ethan. Nothing is worth doing without you guys. The four of you are amazing and have been my biggest cheerleaders. For that, I cannot thank you enough.

I can do all things through Christ who strengthens me. –Philippians 4:13

DEDICATION

I dedicate this dissertation to my husband, Michael Tate. Without your never-ending love, faith, support, and patience, I could not have completed this program. Thank you for your hours of work supporting every other aspect of our life while I worked on courses and research. Thank you for listening to me when things seemed like they couldn't get worse, celebrating when things went right (regardless of how big or little it was), and for being the rational one (yes, you). You are my biggest cheerleader, my best friend, my rock, my love. You are truly my match and I am blessed to have you as my partner. Thank you.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	4
DEDICATION	5
TABLE OF CONTENTS	6
LIST OF FIGURES	15
LIST OF TABLES	16
ABSTRACT	19
Background	19
Methods	19
Results	20
Conclusion	21
ABBREVIATIONS	22
CHAPTER 1. INTRODUCTION	24
1.1 Current Issues and Project Rationale	24
1.2 Project Purpose	24
1.3 Background	25
1.3.1 Introduction	25
1.3.2 Cancer Focus Rationale	26
1.3.3 Clinical Trial Activation	29
1.3.4 Clinical Trial Accrual	31
1.3.5 Factors Affecting Clinical Trial Accrual	33
1.3.6 Site Clinical Trial Infrastructure	35
1.4 Gaps in Literature	37
1.5 Previous Work	38
1.6 Proposed Frameworks	39
1.7 Policy Implications	42
1.8 Specific Aims	42
1.8.1 Specific Aim 1	43
1.8.2 Specific Aim 2	44
1.8.3 Specific Aim 3	46
CHAPTER 2. LITERATURE REVIEW	49
2.1 Introduction to the Issue and Roadmap	49
2.2 Clinical Research Policy Issues	52
2.2.1 A Call for Reform	56
2.3 Clinical Trial Economics and Workload	58
2.4 Study Activation	61
2.4.1 Stage-Gate Framework	64
2.5 Cancer Clinical Research	68

2.5.1 Clinical Trial Pipeline at an NCI-Designated Cancer Center	69
2.5.1.1 Multiple Team Membership (MTM)	71
2.5.2 Scientific Review within the NCI-Designated Cancer Center System	71
2.5.2.1 Feasibility Committees	73
2.6 Clinical Trial Accrual	74
2.6.1 Factors Affecting Clinical Trial Accrual	77
2.6.1.1 Protocol Issues	80
2.6.1.2 Site Issues.....	84
2.6.1.3 Investigator Issues.....	88
2.6.2 Modeling Clinical Trial Accrual.....	91
2.6.2.1 Previous Accrual Prediction Work	94
2.7 Gaps in the Literature.....	95
2.8 Summary	99
CHAPTER 3. METHODOLOGY	101
3.1 Study Purpose	101
3.2 Study Design.....	101
3.2.1 Disease Teams	101
3.3 Center Recruitment	102
3.3.1 Inclusion Criteria	102
3.3.2 Recruitment Methods.....	103
3.4 Data Collection	104
3.4.1 Cancer Center Protocol Data Collection.....	104
3.4.2 Cancer Center Administrative Data Collection	105
3.4.3 Data Integrity	106
3.4.3.1 Data Cleaning.....	106
3.4.3.2 Variable Definitions and Manipulations.....	106
3.4.3.3 Calculation of Protocol per Staffing Unit	113
3.5 Data Analysis	113
3.5.1 General Statistical Considerations	113
3.5.2 Regression Analyses	114
3.5.2.1 Distributions.....	114
3.5.2.2 General Regression Analysis Considerations	116
3.5.2.3 Reduced Models.....	117
3.5.2.4 Full Models	117
3.5.3 Protocol Workload Sensitivity Analyses	118
3.5.4 Accrual Assessments	119
3.6 Assumptions.....	120
3.7 Anticipated Limitations	120
3.8 Computational Support	121
3.9 Confidentiality and Human Subjects	121
CHAPTER 4. RESULTS	123
4.1 Center Recruitment	123

4.1.1 Infrastructure Data Collection.....	127
4.2 Summary Statistics.....	128
4.2.1 Institutional Summary Statistics	128
4.2.1.1 Representativeness of Study Population.....	132
4.2.2 Disease Team Summary Statistics	135
4.2.3 Protocol Summary Statistics	138
4.2.3.1 Zero-Accruing Protocol Summary Statistics	144
4.3 General Statistical Considerations	145
4.4 Specific Aim One.....	146
4.4.1 Contributing Protocol Summary Statistics.....	146
4.4.2 Univariate Analyses and Reduced Models	152
4.4.3 Multivariate Analyses	158
4.4.4 Site-Specific Analysis.....	165
4.4.5 Specific Aim One Summary	169
4.5 Specific Aim Two.....	169
4.5.1 Institutional Infrastructure	169
4.5.2 Contributing Protocols.....	170
4.5.3 Protocol Workload.....	173
4.5.4 Regression Analyses	185
4.5.5 Optimal Protocol Workload for Clinical Trial Accrual	213
4.5.6 Optimal Protocol Workload for Zero-Accruing Protocols	217
4.5.7 Optimal Protocol Workload for Four-Plus Accruing Studies.....	218
4.5.8 Optimal Protocol Workload for Activation Timeline.....	219
4.5.9 Site-Specific Analyses	220
4.5.9.1 Site-Specific Analysis: Protocol Workload	221
4.5.10 Specific Aim Two Summary	233
4.6 Specific Aim Three	234
4.6.1 Univariate Analyses	234
4.6.2 Effect Modifiers and Interactions	237
4.6.3 Full Prediction Model Results – Site-Specific.....	238
4.6.4 Full Prediction Model Results – All Sites.....	244
4.6.5 Model Sensitivity, Specificity, and Accuracy.....	253
4.6.6 Specific Aim Three Summary	257
4.7 Site-Specific Summaries.....	259
4.7.1 Site 104647	259
4.7.1.1 Accrual Prediction	259
4.7.2: Site 173472	261
4.7.2.1 Accrual Prediction	261
4.7.3: Site 448155	263
4.7.3.1 Protocol Workload.....	263
4.7.3.2 Accrual Prediction	265
4.7.4: Site 494048	266
4.7.4.1 Protocol Workload.....	267
4.7.4.2 Accrual Prediction	268

4.7.5: Site 512786	269
4.7.5.1 Protocol Workload	270
4.7.5.2 Accrual Prediction	271
4.7.6: Site 560623	273
4.7.6.1 Protocol Workload	273
4.7.6.2 Accrual Prediction	274
4.7.7: Site 575415	276
4.7.7.1 Accrual Prediction	276
4.7.8: Site 598430	278
4.7.8.1 Accrual Prediction	278
4.7.9: Site 602591	279
4.7.9.1 Protocol Workload	280
4.7.9.2 Accrual Prediction	281
4.7.10: Site 689326	283
4.7.10.1 Accrual Prediction	283
4.7.11: Site 696337	285
4.7.11.1 Protocol Workload	286
4.7.11.2 Accrual Prediction	287
4.7.12: Site 714415	288
4.7.12.1 Protocol Workload	289
4.7.12.2 Accrual Prediction	290
4.7.13: Site 715532	292
4.7.13.1 Accrual Prediction	292
4.7.14: Site 846594	294
4.7.14.1 Accrual Prediction	294
4.7.15: Site 997056	296
4.7.15.1 Accrual Prediction	296
4.7.16: Site 998666	298
4.7.16.1 Accrual Prediction	298
CHAPTER 5. DISCUSSION	300
5.1 Major Findings Related to Clinical Research Site Efficiency	300
5.2 Context of the Problem and the Project	301
5.3 Implications of Findings	302
5.3.1 Specific Aim 1	302
5.3.2 Specific Aim 2	304
5.3.2.1 Considerations for Future Work	306
5.3.4 Specific Aim 3	307
5.3.5 Rare Disease and Pediatric Studies	311
5.3.6 Characteristics of Zero-Accruing Protocols	313
5.3.7 Additional Potential Associated Factors	314
5.4 Limitations	315
5.5 Policy Implications	317
5.6 Conclusions	318

APPENDIX A – EMAIL RECRUITMENT LETTERS.....	319
APPENDIX B – INTRODUCTION LETTER AND ABSTRACTION SHEET.....	322
APPENDIX C – CENTER INFRASTRUCTURE DATA COLLECTION SHEET.....	330
APPENDIX D – FINAL DATABASE DATA DICTIONARY.....	335
APPENDIX E: SITE-LEVEL STUDY CHARACTERISTIC DISTRIBUTION CHARTS	348
Appendix E-1: Distribution of Activation Time in Days.....	348
Appendix E-1a: Site 173472.....	348
Appendix E-1b: Site 448155.....	348
Appendix E-1c: Site 494048.....	349
Appendix E-1d: Site 512786.....	349
Appendix E-1e: Site 560623.....	350
Appendix E-1f: Site 602591	350
Appendix E-1g: Site 696337.....	351
Appendix E-1h: Site 714145.....	351
Appendix E-2: Distribution of Actual Accrual.....	352
Appendix E-2a: Site 104647	352
Appendix E-2b: Site 173472.....	352
Appendix E-2c: Site 448155.....	353
Appendix E-2d: Site 494048.....	353
Appendix E-2e: Site 512786.....	354
Appendix E-2f: Site 560623	354
Appendix E-2g: Site 575415.....	355
Appendix E-2h: Site 598430.....	355
Appendix E-2i: Site 602591.....	356
Appendix E-2j: Site 689326.....	356
Appendix E-2k: Site 696337.....	357
Appendix E-2l: Site 714145.....	357
Appendix E-2m: Site 715532.....	358
Appendix E-2n: Site 846594.....	358
Appendix E-2o: Site 997056.....	359
Appendix E-2p: Site 998666.....	359
APPENDIX F: SPECIFIC AIM TWO – SITE-SPECIFIC PROTOCOL WORKLOAD MODELS	360
Appendix F-1: Protocols per Full-Time Equivalent (FTE) versus Time to Study Activation	360
Appendix F-1a: Site 448155	360
Appendix F-1b: Site 494048.....	360
Appendix F-1c: Site 512786	361
Appendix F-1d: Site 560623.....	361
Appendix F-1e: Site 602591	362
Appendix F-1f: Site 696337.....	362
Appendix F-1g: Site 714145.....	363

Appendix F-2: Protocols per Full-Time Equivalent versus Overall Protocol Accrual	364
Appendix F-2a: Site 448155	364
Appendix F-2b: Site 494048	364
Appendix F-2c: Site 512786	365
Appendix F-2d: Site 560623	365
Appendix F-2e: Site 602591	366
Appendix F-2f: Site 696337	366
Appendix F-2g: Site 714145	367
Appendix F-2h: Site 714145 – Under 100 Subjects	367
Appendix F-3: Protocols per Full-Time Equivalent by Zero-Accruing Status	368
Appendix F-3a: Site 448155	368
Appendix F-3b: Site 494048	368
Appendix F-3c: Site 512786	369
Appendix F-3d: Site 560623	369
Appendix F-3e: Site 602591	370
Appendix F-3f: Site 696337	370
Appendix F-3g: Site 714145	371
Appendix F-4: Protocols per Staff Member versus Time to Study Activation	372
Appendix F-4a: Site 448155	372
Appendix F-4b: Site 494048	372
Appendix F-4c: Site 512786	373
Appendix F-4d: Site 560623	373
Appendix F-4e: Site 602591	374
Appendix F-4f: Site 696337	374
Appendix F-4g: Site 714145	375
Appendix F-5: Protocols per Staff Member versus Overall Protocol Accrual	376
Appendix F-5a: Site 448155	376
Appendix F-5b: Site 494048	376
Appendix F-5c: Site 512786	377
Appendix F-5d: Site 560623	377
Appendix F-5e: Site 602591	378
Appendix F-5f: Site 696337	378
Appendix F-5g: Site 714145	379
Appendix F-5h: Site 714145 – Under 100 Subjects	379
Appendix F-6: Protocols per Staff Member by Zero-Accruing Status	380
Appendix F-6a: Site 448155	380
Appendix F-6b: Site 494048	380
Appendix F-6c: Site 512786	381
Appendix F-6d: Site 560623	381
Appendix F-6e: Site 602591	382
Appendix F-6f: Site 696337	382
Appendix F-6g: Site 714145	383
Appendix F-7: Efficiency Outcomes by Protocol Workload	384
Appendix F-7a: Site 448155 – Protocols per Staff Member	384

Appendix F-7b: Site 448155 – Protocols per FTE.....	384
Appendix F-7c: Site 494048 – Protocols per Staff Member/FTE	385
Appendix F-7d: Site 512786 – Protocols per Staff Member/FTE	386
Appendix F-7e: Site 560623 – Protocols per Staff Member/FTE	386
Appendix F-7f: Site 602591 – Protocols per Staff Member	387
Appendix F-7g: Site 602591 – Protocols per FTE.....	388
Appendix F-7h: Site 696337 – Protocols per Staff Member/FTE	389
Appendix F-7i: Site 714145 – Protocols per Staff Member/FTE.....	389
Appendix F-8 Site-Specific Full Models Assessing the Impact of Protocol Workload on Clinical Trial Accrual	391
Appendix F-8a: Site 494048 – Protocols per Staff Member and Protocols per FTE	391
Appendix F-8b: Site 560623 – Protocols per Staff Member and Protocols per FTE	393
Appendix F-8c: Site 602591 – Protocols per Staff Member.....	394
Appendix F-8d: Site 602591 – Protocols per FTE.....	396
Appendix F-8e: Site 714145 – Protocols per Staff Member and Protocols per FTE	398
Appendix F-9 Site-Specific Full Models Assessing the Impact of Protocol Workload on Zero- Accruing Trials	401
Appendix F-9a: Site 494048 – Protocols per Staff Member and Protocols per FTE	401
Appendix F-9b: Site 560623 – Protocols per Staff Member and Protocols per FTE	403
Appendix F-9c: Site 602591 – Protocols per Staff Member.....	405
Appendix F-9d: Site 602591 – Protocols per FTE.....	407
Appendix F-10 Site-Specific Full Models Assessing the Impact of Protocol Workload on the Likelihood of a Trial Accruing at Least Four Subjects.....	410
Appendix F-10a: Site 494048 – Protocols per Staff Member and Protocols per FTE	410
Appendix F-10b: Site 512786 – Protocols per Staff Member and Protocols per FTE	412
Appendix F-11 Site-Specific Full Models Assessing the Impact of Protocol Workload on Activation Time	415
Appendix F-11a: Site 494048 – Protocols per Staff Member and Protocols per FTE	415
Appendix F-11b: Site 602591 – Protocols per Staff Member	417
Appendix F-11c: Site 602591 – Protocols per FTE.....	419
Appendix F-11d: Site 696337 – Protocols per Staff Member and Protocols per FTE	421
APPENDIX G: SPECIFIC AIM THREE – SITE MODELS PREDICTING CLINICAL TRIAL ACCRUAL	423
Appendix G-1: Site-Specific Models Predicting Clinical Trial Accrual	423
Appendix G-1a: Site 104647	423
Appendix G-1b: Site 173472	424
Appendix G-1c: Site 448155	426
Appendix G-1d: Site 494048	428
Appendix G-1e: Site 512786	430
Appendix G-1f: Site 560623	431
Appendix G-1g: Site 575415	433
Appendix G-1h: Site 598430	434
Appendix G-1i: Site 602591	436

Appendix G-1j: Site 689326	437
Appendix G-1k: Site 696337	439
Appendix G-1l: Site 714415	440
Appendix G-1m: Site 715532	442
Appendix G-1n: Site 846594	443
Appendix G-1o: Site 997056	445
Appendix G-1p: Site 998666	446
Appendix G-2: Site-Specific Prediction Model versus Actual Accrual	448
Appendix G-2a: Site 104647	448
Appendix G-2b: Site 173472	448
Appendix G-2c: Site 448155	449
Appendix G-2d: Site 494048	449
Appendix G-2e: Site 512786	450
Appendix G-2f: Site 560623	450
Appendix G-2g: Site 575415	451
Appendix G-2h: Site 598430	451
Appendix G-2i: Site 602591	452
Appendix G-2j: Site 689326	452
Appendix G-2k: Site 696337	453
Appendix G-2l: Site 714415	453
Appendix G-2m: Site 715532	454
Appendix G-2n: Site 846594	454
Appendix G-2o: Site 997056	455
Appendix G-2p: Site 998666	455
Appendix G-3: Disease Team Predicted versus Actual Accrual	456
Appendix G-3a: Site 104647	456
Appendix G-3b: Site 173472	456
Appendix G-3c: Site 448155	457
Appendix G-3d: Site 494048	457
Appendix G-3e: Site 512786	458
Appendix G-3f: Site 560623	458
Appendix G-3g: Site 575415	459
Appendix G-3h: Site 598430	459
Appendix G-3i: Site 602591	460
Appendix G-3j: Site 689326	460
Appendix G-3k: Site 696337	461
Appendix G-3l: Site 714415	461
Appendix G-3m: Site 715532	462
Appendix G-3n: Site 846594	462
Appendix G-3o: Site 997056	463
Appendix G-3p: Site 998666	463
Appendix G-4: Overall Prediction Model versus Actual Accrual	464
Appendix G-4a: Site 104647	464
Appendix G-4b: Site 173472	464

Appendix G-4c: Site 448155	465
Appendix G-4d: Site 494048	465
Appendix G-4e: Site 512786	466
Appendix G-4f: Site 560623	466
Appendix G-4g: Site 575415	467
Appendix G-4h: Site 598430	467
Appendix G-4i: Site 602591	468
Appendix G-4j: Site 689326	468
Appendix G-4k: Site 696337	469
Appendix G-4l: Site 714415	469
Appendix G-4m: Site 715532	470
Appendix G-4n: Site 846594	470
Appendix G-4o: Site 997056	471
Appendix G-4p: Site 998666	471
Appendix G-5: Overall Prediction Model: Random Effects versus Actual Accrual	472
Appendix G-5a: Site 104647	472
Appendix G-5b: Site 173472	472
Appendix G-5c: Site 448155	473
Appendix G-5d: Site 494048	473
Appendix G-5e: Site 512786	474
Appendix G-5f: Site 560623	474
Appendix G-5g: Site 575415	475
Appendix G-5h: Site 598430	475
Appendix G-5i: Site 602591	476
Appendix G-5j: Site 689326	476
Appendix G-5k: Site 696337	477
Appendix G-5l: Site 714415	477
Appendix G-5m: Site 715532	478
Appendix G-5n: Site 846594	478
Appendix G-5o: Site 997056	479
Appendix G-5p: Site 998666	479
REFERENCES	480

LIST OF FIGURES

Figure 1-1: Factors Affecting Clinical Trial Accrual	34
Figure 2-2: Simplified Study Activation Pipeline	72
Figure 1-1: Factors Affecting Clinical Trial Accrual	78
Figure 2-1: Factors Affecting Clinical Trial Accrual Considering Time	79
Figure 2-3: Key Clinical Trial Players.....	96
Figure 3-1: Continuous Variable Transformation Matrix.....	111
Figure 3-1a: Months of Accrual Completed	111
Figure 3-1b: National Enrollment Goal	111
Figure 3-1c: Total Months of Accrual	112
Figure 3-1d: Activation Time	112
Figure 4-1: Participating Centers	126
Figure 4-2 All Centers with Participating Centers Highlighted.....	133
Figure 4-3: Distribution of Activation Time – All Sites.....	143
Figure 4-4: Distribution of Actual Accrual – All Sites.....	144
Figure 4-5: Protocols per Full-Time Equivalent versus Overall Protocol Accrual	175
Figure 4-5a: All Sites	175
Figure 4-5b: All Sites: Accrual <100.....	175
Figure 4-6: Protocols per Staff Member versus Overall Protocol Accrual – All Sites.....	176
Figure 4-6a: All Sites	176
Figure 4-6b: All Sites: Accrual <100.....	176
Figure 4-7: Protocol Workload versus Time to Study Activation – All Sites	177
Figure 4-7a: Protocols per FTE.....	177
Figure 4-7b: Protocols per Staff Member	177
Figure 4-8 Mean Clinical Trial Accrual by Protocol Workload.....	178
Figure 4-8a: Protocols per Staff Member	178
Figure 4-8b: Protocols per FTE	178
Figure 4-9 Mean Activation Time by Protocol Workload.....	179
Figure 4-9a: Protocols per Staff Member	179
Figure 4-9b: Protocols per FTE	179

LIST OF TABLES

Table 3-1: Dependent-Independent Variable Pairings.....	115
Table 4-1: Institution Summary Statistics.....	129
Table 4-2: Study Accrual Summary Statistics	130
Table 4-3: Disease Team Prediction Summary Statistics	136
Table 4-4: Protocol Summary Statistics by Accrual Status	138
Table 4-5: Feasibility Committee, Regulatory Staffing Model, and Study Activation Information	147
Table 4-6: Summary Statistics by Stage-Gate Classification	148
Table 4-7: Adjusted, Univariate Models Assessing the Impact of the 5-stage/3-gate versus the 3-stage/2-gate Framework.....	153
Table 4-7a: Stage-Gate Process Impact on Clinical Trial Accrual	153
Table 4-7b: Stage-Gate Process Impact on the Likelihood of Zero-Accruing Studies.....	154
Table 4-7c: Stage-Gate Process Impact on the Likelihood of Accruing at Least Four Subjects	154
Table 4-7d: Stage-Gate Process Impact on Activation Timeline.....	155
Table 4-8: Variable Association with Clinical Trial Accrual by Stage-Gate Process	156
Table 4-9: Effect Modification and Interaction Analyses by Stage-Gate Process.....	157
Table 4-10: Variable Association with Activation Time* by Stage-Gate Process	158
Table 4-11: Full Model Assessing the Impact of Stage-Gate Process on Clinical Trial Accrual.....	160
Table 4-9: Effect Modification and Interaction Analyses by Stage-Gate Process.....	162
Table 4-12: Full Model Assessing the Impact of Stage-Gate Process on Activation Time	163
Table 4-13: Single Site Analysis: 696337 Univariate Models Assessing the Impact of the 5-stage/3-gate versus the 3-stage/2-gate Framework.....	166
Table 4-13a: Stage-Gate Process Impact on Clinical Trial Accrual	166
Table 4-13b: Stage-Gate Process Impact on the Likelihood of Zero-Accruing Studies.....	166
Table 4-13c: Stage-Gate Process Impact on the Likelihood of Accruing at Least Four Subjects	166
Table 4-13d: Stage-Gate Process Impact on Activation Timeline.....	167
Table 4-14: Single Site: Variable Association with Activation Time* by Stage-Gate Process .	167
Table 4-15: Single Site Analysis: 696337 Full Model Assessing the Impact of Stage-Gate Process on Activation Time	168
Table 4-16: Summary Statistics for Institutions Participating in Specific Aim Two	171
Table 4-17: Protocol Workload Summary Statistics by Site	174

Table 4-18: Efficiency Outcomes by Protocol Workload (Protocols per Staff Member) – Participating Sites	181
Table 4-19: Efficiency Outcomes by Protocol Workload (Protocols per FTE) – Participating Sites.....	183
Table 4-20: Univariate Adjusted Models Assessing Protocol Workload Impact – All Sites	186
Table 4-20a: Protocol Workload (Protocols per Staff Member) on Clinical Trial Accrual ...	186
Table 4-20b: Protocol Workload (Protocols per Staff Member) on the Likelihood of Zero-Accruing Studies	187
Table 4-20c: Protocol Workload (Protocols per Staff Member) on the Likelihood of Accruing at Least Four Subjects	188
Table 4-20d: Protocol Workload (Protocols per Staff Member) on Activation Timeline	189
Table 4-20e: Protocol Workload (Protocols per FTE) on Clinical Trial Accrual.....	190
Table 4-20f: Protocol Workload (Protocols per FTE) on the Likelihood of Zero-Accruing Studies.....	191
Table 4-20g: Protocol Workload (Protocols per FTE) on the Likelihood of Accruing at Least Four Subjects	192
Table 4-20h: Protocol Workload (Protocols per FTE) on Activation Timeline	193
Table 4-21: Variable Association with Clinical Trial Accrual by Protocol Workload – All Sites	194
Table 4-22: Effect Modification and Interaction Analyses by Protocol Workload – All Sites ..	195
Table 4-23: Full, Adjusted Model Assessing the Impact of Protocols per Staff Member on Clinical Trial Accrual	197
Table 4-24: Full, Adjusted Model Assessing the Impact of Protocols per Full-Time Equivalent (FTE) on Clinical Trial Accrual.....	199
Table 4-25: Variable Association with Activation Time* by Protocol Workload – All Sites ...	202
Table 4-26: Full, Adjusted Model Assessing the Impact of Protocols per Staff Member on Activation Time	204
Table 4-27: Full, Adjusted Model Assessing Protocols per Full-Time Equivalent (FTE) on Activation Time	206
Table 4-28: Variable Association with Likelihood to Accrue at Least Four Subjects by Protocol Workload – All Sites.....	209
Table 4-29: Full, Adjusted Model Assessing Protocols per Staff Member on Likelihood of Accruing 4+ Subjects.....	211
Table 4-30: Best Workload Performance by Study Characteristic Analysis – Protocols per Staff Member	214
Table 4-31: Best Workload Performance by Study Characteristic Analysis – Protocols per FTE	215

Table 4-32: Site-Specific Protocol Workload Univariate Analyses	221
Table 4-33: Independent Variable Association of Protocol Workload with Clinical Trial Accrual by Site	223
Table 4-34: Independent Variable Association of Protocol Workload with Zero-Accruing Protocols by Site	225
Table 4-35: Independent Variable Association of Protocol Workload with At Least Four Accruals by Site	227
Table 4-36: Independent Variable Association of Protocol Workload with Activation Time by Site	228
Table 4-37: Effect Modification and Interaction Analyses for Protocol Workload by Site	230
Table 4-38: Optimal Protocol Workload by Regression Analysis.....	232
Table 4-39: Independent Variable Association with Clinical Trial Accrual by Site	235
Table 4-40: Effect Modification and Interaction Analyses with Clinical Trial Accrual by Site	237
Table 4-41: Actual and Predicted Accrual Models by Site.....	240
Table 4-42: Correlation of Predicted Accrual to Actual Accrual	241
Table 4-43: Site-Specific Prediction Model versus Actual Accrual – All Sites	242
Table 4-44: Correlation of Categories between Predicted Accrual and Actual Accrual	243
Table 4-45: Disease Team Predicted versus Actual Accrual – All Sites.....	244
Table 4-46: Independent Variable Association with Clinical Trial Accrual for All Sites.....	245
Table 4-47: Effect Modification and Interaction Analyses with Clinical Trial Accrual for All Sites.....	246
Table 4-48: Full, Unadjusted Model Predicting Clinical Trial Accrual for All Sites.....	248
Table 4-49: Full, Adjusted Random Effects Model Predicting Clinical Trial Accrual for All Sites	249
Table 4-50: Overall Prediction Model versus Actual Accrual – All Sites.....	252
Table 4-51: Overall Prediction Model: Random Effects versus Actual Accrual – All Sites.....	252
Table 4-52: Sensitivity, Specificity, and Accuracy for Accrual Prediction Models using a Four Accrual Cut-off	254
Table 4-53: Best Accrual Prediction Model by Site	258

ABSTRACT

Background

Clinical trials represent a significant percentage of the time and cost to bring a drug through the development process and to Food and Drug Administration approval. Despite how critical these trials are to the drug development process, many studies are underpowered due to low accrual. This translates to valuable questions regarding the safety and effectiveness of new agents being left unanswered, requiring additional time and studies. A call for reform of the industry has been made by stakeholders in the clinical research enterprise; however, national change is slow. Thus sites that conduct clinical research must find methods to increase efficiency within the burdensome system currently in place. Throughout cancer centers adhering to the National Cancer Institute (NCI) Cancer Center Support Grant guidelines, efficiencies have been explored individually; however, there is a gap in knowledge on what factors affect sites system-wide. This dissertation seeks to examine factors that affect clinical trial efficiency in the areas of study activation looking at the outcome of local clinical trial accrual.

Methods

Protocol and site-specific clinical trial administration data was collected regarding closed, interventional treatment and supportive care clinical trials from cancer centers adhering to NCI Cancer Center Support Grant guidelines during a five-year time period (2009-2014). Study characteristic analyses and hierarchical regression modeling was used to explore the effect of feasibility committee use and protocol workload on the outcomes of clinical trial accrual and time to activate a clinical trial. Sensitivity analyses were utilized when considering protocol

workload to account for studies that had not yet closed to accrual, and thus were not included in this dataset. In addition, protocol- and site-specific variables were used to build regression models used to predict clinical trial accrual. Sensitivity, specificity, and accuracy were compared to the current standard, the institutional disease team.

Results

Sixteen centers contributed a total of 5,787 protocols (range 93-697 studies). These studies accrued 49,319 subjects. Of all studies, 1,053 (18%) accrued zero subjects. Disease teams predicted 221% of actual accrual. Seven institutions submitted protocol workload information for 2,133 studies (36.9%) and 14,229 accruals (28.9%). Controlling for effect modifiers and interactions, and adjusting for institution, a statistically significant increase in clinical trial accrual and decrease in activation time was seen with the use of a feasibility committee. Regulatory protocol workload was significantly associated with clinical trial accrual and activation time; however, a single, definitive protocol workload was not identified that both minimized activation time and maximized clinical trial accrual. Protocol workload most often maximized accrual at workloads of between 3.5 and 5.0 protocols per staff member/FTE and minimized activation time at workloads between 1.0 and 1.9 protocols per staff member/FTE. Regression models predicted accrual more accurately than disease teams at all 16 centers, with site-specific models consistently having the best performance (versus an adjusted, hierarchical model).

Conclusion

Despite institutional differences in variable association with accrual and activation times, the utilization of a feasibility committee was shown to improve clinical trial accrual as well as decrease activation time. Using systematic methods for examining study activation and accrual efficiencies resulted in the development of models that predicted clinical trial accrual better than the current standard (disease team prediction) at all participating centers. Further research is needed to better define and determine optimal workload. This information and these models may better inform study planning and resource allocation decisions by local stakeholders (administrators and investigators) in the clinical research enterprise.

ABBREVIATIONS

AACI-CRI: American Association of Cancer Institutes Clinical Research Initiative

ACS: American Cancer Society

AIC: Akaike Information Criterion

ASCO: American Society of Clinical Oncology

BCPT: Breast Cancer Prevention Trial

BIC: Bayesian Information Criterion

CALGB: Cancer and Leukemia Group B

CAST: Cardiac Arrhythmia Suppression Trial

CCOP: Community Clinical Oncology Program

CCSG: Cancer Center Support Grant

CPDM: Clinical Protocol and Data Monitoring

CSDD: Center for the Study of Drug Development

CTE: Clinical trial enterprise

CTEP: Cancer Therapy Evaluation Program

CTMS: Clinical trial management system

CTO: Clinical trials office

CTSA: Clinical and Translational Science Award

DHHS: Department of Health and Human Services

ECOG: Eastern Cooperative Oncology Group

FDA: Food and Drug Administration

FTE: Full-time equivalent

HIV/AIDS: Human Immunodeficiency Virus/Acquired Immune Deficiency Syndrome

I-O: Industrial-organization

ICMJE: International Committee of Medical Journal Editors

IOM: Institute of Medicine

IRB: Institutional Review Board

LOI: Letter of intent

MTM: Multiple Team Membership

NAAL: National Assessment of Adult Literacy

NBER: National Bureau of Economic Research

NCI: National Cancer Institute

NCTN: National Clinical Trials Network

NIH: National Institutes of Health

PI: Principal investigator

POSCH: Program on Surgical Control of the Hyperlipidemias

PRMS: Protocol Review and Monitoring System

R&D: Research and development

RCT: Randomized controlled trial

REACT: Recruitment and Enrollment Assessment in Clinical Trials

RVU: Relative value unit

SWE: Site work effort

UACC: University of Arizona Cancer Center

U.S.: United States

CHAPTER 1. INTRODUCTION

“We must all hang together, or assuredly we shall all hang separately.”

-Benjamin Franklin

1.1 Current Issues and Project Rationale

The drug development process, specifically clinical trials, is a lengthy and expensive process, with time and cost factors increasing annually. Several entities within the United States (U.S.) as well as internationally have called for reform in the clinical trial enterprise; however, such change is slow (1-3). Lasting change will require federal intervention and modification of the national research infrastructure. While the need for change is recognized by the federal government and research agencies, clinical trial study designs and technologies are evolving at a rapid rate. This leaves institutions conducting clinical research in a lurch as they must execute clinical trials utilizing these new technologies while confined to the outdated constraints of the regulations, guidance, and requirements of sponsors, regulatory agencies, and federal oversight agencies. Thus effective measures for conducting innovative clinical research within the current paradigm at the site level must be developed, which is the focus of this dissertation project.

1.2 Project Purpose

The purpose of this dissertation project is to address the gap in knowledge regarding how to objectively describe and quantitate optimal study activation workload at the study site level while maximizing the probability of clinical trial success, defined as site clinical trial accrual. This will be done utilizing existing industrial-organization (I-O) framework applied to the

clinical research enterprise of cancer centers adhering to National Cancer Institute (NCI) guidelines. This evaluation of the clinical trial activation process will use two I-O frameworks. The first, stage-gate framework, examines the integration of formal review processes to increase the proportion of successful project launches, defined here as clinical trial protocols that accrue subjects. The second, Multiple Team Membership (MTM) framework, correlates the number of simultaneous projects a staff member has with productivity, defined here as a reduction in the time it takes to activate a study. In addition, to provide quantitative information regarding the likelihood that a study will succeed, predictive regression modeling—a statistical method to measure the association between two variables while controlling for other contributing factors—will be utilized with the outcome of clinical trial accrual. These three processes will assist in the development of a multi-faceted evaluation process with quantitative evidence to support efficient decision-making regarding which trials to activate and needed staffing levels.

1.3 Background

1.3.1 Introduction

Dr. Harold Varmus, past-Director of the National Institutes of Health (NIH), U.S. Department of Health and Human Services (DHHS), and past-Director of the NCI, stated to the Senate Committee on Labor and Human Resources, Subcommittee on Public Health and Safety, on October 9, 1997, that “clinical research has changed the face of modern medicine” (4). Dr. Varmus’ statement regarding the clinical research enterprise was corroborated by the Institute of Medicine (IOM) in 2010 when they stated that “[e]fficiently generating medical evidence and translating it into practice implies a ‘learning health care system’ in which the divide between

clinical practice and research is diminished and ultimately eliminated” (1). However, this process of utilizing clinical research to change modern medicine is a slow and meticulous one. The Tufts Center for the Study of Drug Development (CSDD) has reported that it takes approximately \$2.6 billion (U.S., 2013) and 128 months of development to bring a drug through pre-clinical development to Food and Drug Administration (FDA) approval, with 80% of compounds being discontinued at some point of the process (5). Of the 128 months of development, 97 of them are spent in clinical testing (*ibid*). The Lancet, in a series of review articles focused on research waste, reported that it takes over 20 years to translate basic science into clinical practice with very few discoveries widely utilized in the clinical realm (6).

During the 10+ years it takes to get a new drug to market, patients needing that treatment are left without, possibly with lethal results. For those who do benefit from the marketing of a new drug, they must face increasing drug costs, which are partially based on the high cost of development. This leads to financial repercussions (including bankruptcy) to the patient and their family. Therefore it is imperative that something be done to remedy this situation in order to better the health of the population at the individual level. If the clinical development of a drug, which takes three-quarters of the time in the drug development process, can be made more efficient, then those savings will be passed along to society through the advancement of medicine.

1.3.2 Cancer Focus Rationale

Despite over \$105 billion spent by the NCI over the past 40+ years and \$3.4 billion spent by the American Cancer Society (ACS) over the past approximately 70 years studying cancer, overall

incidence and mortality rates in many cancer subtypes are not improving (7). Cancer is the second largest cause of death in the U.S., attributed to 584,881 deaths in 2013 (8). It is projected to surpass heart disease as the number one killer in the U.S. by 2030 (9). According to a report of industry costs in drug development, oncology is the third most expensive area of clinical trial research, with the estimated cost of the clinical development of one drug exceeding \$78 million (U.S. 2008, assuming one study per phase, phases I-III; 10). A search of ClinicalTrials.gov on June 23, 2015 showed 5,584 open, interventional cancer (conditions = cancer OR neoplasm OR oncology OR carcinoma) studies (not yet recruiting, recruiting, or expanded access) in the U.S. This means that almost one-third (32.7%) of all open, interventional clinical trials in the U.S. are focused on the study of cancer.

With the advent and implementation of genome-wide testing, cancer tumors are found to contain unique molecular signatures, differing from normal cells, other cancers of the same anatomical origin, and even from other cancer cells within the same tumor. Because of this cancer is actually hundreds if not thousands of different diseases. Highlighting this fact, the NCI defines cancer as “the name given to a *collection* of related diseases” (emphasis added; 11). The ongoing discoveries of these different genetic drivers of cancer calls for increased research to determine which are the most effective and safest treatments. This is at the heart of President Obama’s and the NIH’s Precision Medicine Initiative (PMI). Announced January 30, 2015, President Obama has invested \$215 million of the 2016 budget to this initiative, which will advance patient-driven treatment based on the genetic uniqueness of their disease. One of the primary objectives of the Precision Medicine Initiative is for the NCI to “accelerate the design

and testing of effective, tailored treatments for cancer...” using \$70 million, or 32.6% of the total budget (12). While other diseases will benefit from the Precision Medicine Initiative, cancer is the near-term focus (13). Continuing this initiative is the “moonshot” mission, announced during President Obama’s State of the Union address on January 12, 2016, and spearheaded by Vice President Biden. This mission has the ambitious goal of curing cancer, and thus conducting large amounts of clinical research, by 2020 (14).

The scope of this dissertation work is cancer centers adhering to the NCI Cancer Center Support Grant (CCSG) guidelines. This population provides a homogenous pool of study sites that conduct a large volume of clinical research with support staff dedicated to conducting clinical research (and not involved in clinical practice). These centers all have clinical trial offices (CTOs) that manage some aspect of clinical research under guidelines outlined by the NCI, providing a base level of standardization. The use of this population will allow conclusions to be drawn regarding the impact of processes and tools on workflow and defined outcomes (accrual and activation) without the potential confounding factors of clinical practice. Additionally, as the pool of cancer protocols is large—stated earlier as approximately one-third of all currently open clinical trials on ClinicalTrials.gov—a large homogenous pool of clinical research protocols can be assessed, reducing the potential for bias due to differential selection (e.g. research in different disease states impacts the outcomes of interest differently due to the nature of their design).

1.3.3 Clinical Trial Activation

On average, it takes 6.3 years to complete clinical investigations (phase I-III testing) for a single agent. While FDA approval for oncology drugs is approximately six months faster than non-oncology drugs, on average oncology drugs spend 1.5 years longer (for an average of 7.8 years total) in the clinical investigation phase of development than non-oncology drugs (15). It is also reported that the length of this process has increased over time (16). An analysis of site work effort showed that, between 1989 and 2011, workload increased approximately three-fold and the number of subjects per site has decreased from a median of 20 to a median of eight. Oncology made up 9% of the protocol contracts in this database (5.89% in 1989, increasing to 12.55% in 2011; 17). Further analyses of the clinical trial pipeline shows that it is not just the overall clinical testing process that is lengthy, but each clinical trial itself. The clinical trial process was mapped out by Dr. David Dilts in the mid-2000s, both at the national, cooperative group level (18-19) and the local cancer center site level (18, 20). In his review of the phase III clinical trial development and activation process within the NCI cooperative group setting, Dilts reported that it took 769 steps, 36 approvals, and a median of 2.5 years from formal concept review to the study opening to accrual (18). Ninety-six (96) of these steps took place at the comprehensive cancer center (site) level (*ibid*). A review of the Cancer and Leukemia Group B (CALGB) cooperative group phase III clinical trial process showed that there were 370 distinct processes needed to open a study, with a median time from initial concept to study activation of 784 days (range: 537-1130 days; 19). In a separate study, Dilts' analysis of the comprehensive cancer center process at Vanderbilt University showed that 50% of the steps to open an oncology clinical trial held no value (20). At the time of the report, it took a median of 171 days to open a

clinical trial at the Vanderbilt cancer center (95% confidence interval: 158-192 days; range: 27-657 days). Since the publishing of these papers, formal and informal review committees have been implemented at several NCI-designated cancer centers to address clinical trial feasibility, adding additional steps to an already lengthy process (personal correspondence).

Based on these observed trends in the clinical trial enterprise (CTE), the IOM held a two-day workshop in 2011 focused on the foundation of the clinical trial enterprise, randomized controlled clinical trials (RCTs; 21). In this report, compiled by 40 experts in the field of clinical research hailing from academia, the federal government, non-profit organizations, and pharmaceutical companies, a call for research business plans was made to speed cultural changes in organization. This recommendation was made due to a “widening separation” between clinical research and clinical practice (the health care system) and noted lagging cultural changes in the research milieu while research technology (e.g. genetic mutation screening, cancer subtyping, and updated study designs) has advanced rapidly (*ibid*). A separate report in the Lancet echoed the sentiment of the IOM, stating that research waste (monetary, human workforce, and time resources) was being created due to 1) increased regulatory processes, 2) inconsistencies within these processes, and 3) the timelines to conduct both basic and clinical (observational and interventional) research (3).

The IOM workshop categorized the clinical trial enterprise into four laboratories:

- 1) **Innovation**, or to develop initial evidence about treatments and biological markers and hypotheses

- 2) **Traditional Clinical Research**, or the determination of treatment efficacy, or risks and benefits in carefully defined populations
- 3) **Health Care Delivery**, the evaluation of treatment risks and benefits in the context of health care
- 4) **Community Engagement**, the assessment of strategies of disease prevention and wellness (including living with chronic disease) (21).

The focus of this dissertation work is on the second category, Traditional Clinical Research.

This section of the clinical trial enterprise encompasses the FDA drug approval process which, as reported by DiMasi and stated earlier in this section, takes 97 of the total 128 months to get a drug from preclinical development through approval (5), representing a significant bottleneck in getting safe and effective drugs to market.

1.3.4 Clinical Trial Accrual

Despite the length of time dedicated to the clinical trial portion of drug approval, many studies do not meet their accrual goals (16). Highlighted by Denny et al. in 2004, 9.4% of CALGB protocols from 1989 to 1997 did not meet accrual goals and ended early, generating too little data to answer the questions they sought to answer (22). Supporting Denny's 2004 report regarding under-accruing studies are several other studies looking at cancer clinical trial accrual. According to a 2010 IOM report, 40% of NCI cooperative group phase III trials opened between 2000 and 2007 did not meet their minimum accrual goal (2). In the same year, Dilts et al. reported that 6.4% of Cancer Therapy Evaluation Program (CTEP), the NCI program that vets NCI-sponsored clinical trials, did not accrue a single subject nationally (18). Schroen et al.

reported that from 1998-2002, 52.5% of phase III trials within a single NCI cooperative group closed due to insufficient accrual, defined as not having met the accrual goal to answer the scientific question (23). Forty-five percent (45%) of trials from a sample of all participating NCI cooperative groups during the same time frame closed with insufficient accrual (*ibid*).

Approximately 38% of CTEP trials open to patient accrual between June 1, 2000, and December 31, 2004, did not attain their accrual goal, with approximately 70% of phase III studies not meeting their predicted accrual goal (24). In addition, Dilts reported that at four model comprehensive cancer centers, it is more likely that an NCI-sponsored cooperative group trial will not accrue any subjects at the site than a study that is not sponsored by the cooperative group (38.8% versus 20.6%; 18). Regardless, both percentages represent a large number of studies that fail to meet pre-determined accrual goals. Another study reported that the mean percent of zero-accruing studies at four comprehensive cancer centers was 28.6% (range 20.6%-34.4%; 25). The IOM 2010 workshop summary reports this statistic at 27% (1) and an American Society of Clinical Oncology (ASCO) abstract reported this number as 50% compiled from 14 cancer centers (26). Not obtaining the needed number of research subjects during the dedicated recruitment period leads to prolonged recruitment phases, which increases the cost of conducting the study. The IOM reports that clinical trial enrollment periods are estimated to be extended in approximately 90% of trials worldwide (1). This leads to an increase in the number of sites and staff needed to conduct clinical research, inflating costs (3). Underpowered studies mean that the scientific question at hand is not properly addressed, leading to research waste (27). This also is an issue of public health, as it delays the availability of safe and effective treatments to the general public (28).

1.3.5 Factors Affecting Clinical Trial Accrual

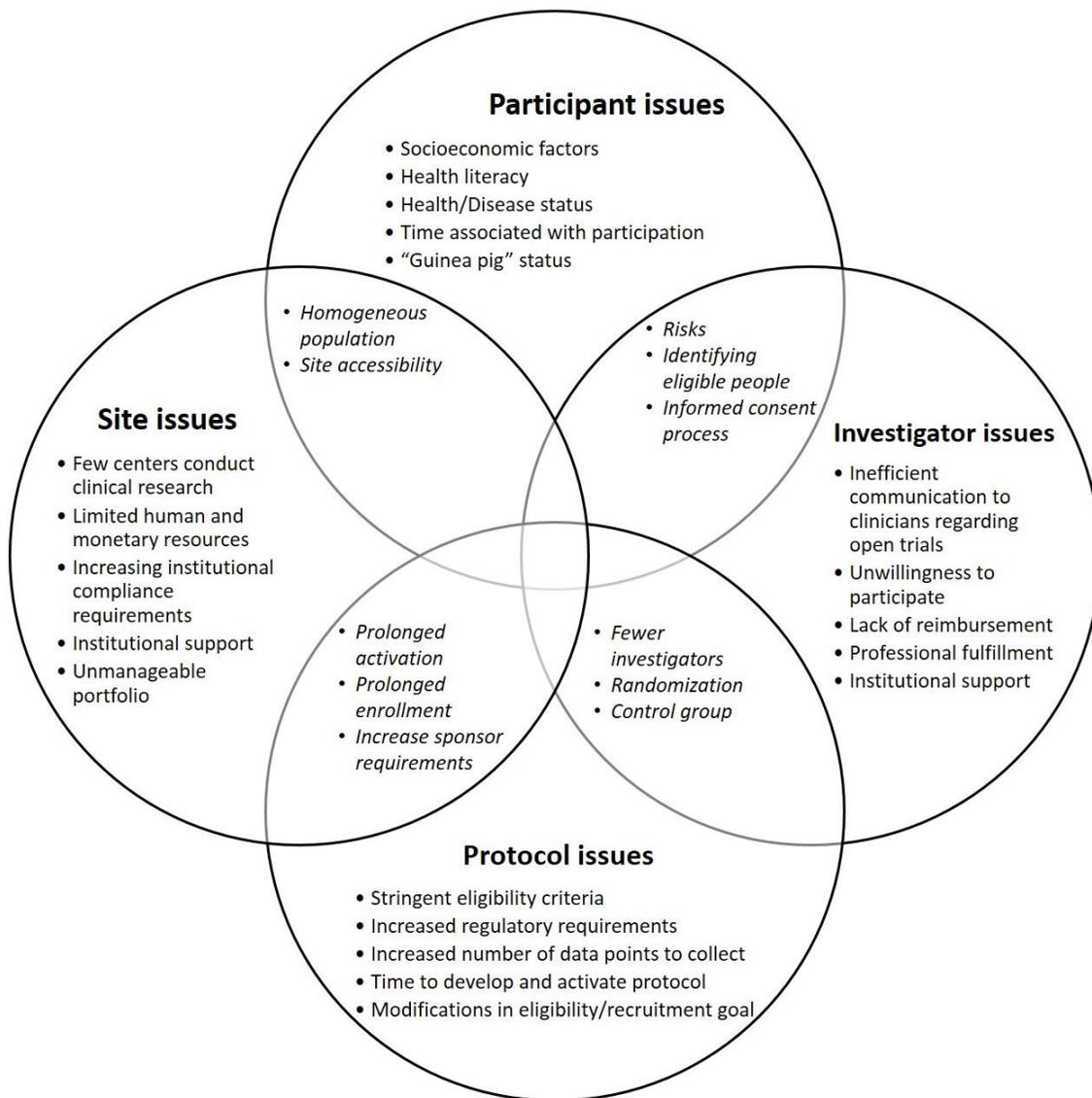
Four general categories that impact clinical trial accrual are diagramed in **Figure 1-1** and include:

- 1) **Protocol issues:** Stringent eligibility criteria (21), increased regulatory requirements (such as conflict of interest and privacy review; 3, 29), increasing protocol complexity (10), and numerous procedures requiring significant participant time (30-31).
- 2) **Site issues:** It is estimated that two-thirds of sites never do more than a single trial (32) and approximately half of clinical trial accruals come from academic centers (33). An analysis of clinical trials show that a large proportion of trials are sponsored by academic health centers focused on disease mechanisms (34). Conducting such studies requires dedicated human resources. Additionally, as clinical trials become more complicated (21, 29), specialized staff and/or training is needed to execute clinical research efficiently and in a compliant manner.
- 3) **Investigator issues:** Inefficiencies in communicating information regarding open clinical research studies to clinicians, unwillingness for clinicians to participate in research, and a lack of reimbursement for research procedures (21). There is a decline in the number of U.S. investigators with an increase in clinical trials worldwide and global outsourcing of clinical trial work (21, 34). The IOM reports that, as of 2007, 85% of investigators only conduct one clinical trial (1).
- 4) **Participant issues:** Due to social factors, financial issues, or simply the ability to identify and find eligible persons to enroll (35-38).

This dissertation focuses primarily on site and protocol issues.

Figure 1-1: Factors Affecting Clinical Trial Accrual

Four non-mutually exclusive categories under which factors affecting clinical trial accrual can be classified.



1.3.6 Site Clinical Trial Infrastructure

Many academic sites are centralizing their clinical trials staff and utilizing NIH program P30 federal funding mechanisms, such as the Cancer Center Support Grant (CCSG) and Clinical and Translational Science Award (CTSA), to provide significant infrastructure support for these offices. These clinical trials offices provide a core of professional staff dedicated to providing clinical trial support in at least one of the following areas: regulatory support and compliance, clinical research coordination, data management, internal monitoring, research nursing, or clinical research informatics.

The Clinical and Translational Science Award program, hosted by the NIH, seeks to address the “development and implementation of national standards and best practices for translation, from basic discovery to clinical and community-engaged research (39).” The group of institutions formed to address these standards is the Clinical and Translational Science Award Consortium. One focus of the Clinical and Translational Science Award Consortium is to tackle clinical trial recruitment roadblocks, citing an inability of trials to completely recruit the needed participants to complete trials (16). A review of the Clinical and Translational Science Award program categorized three barriers to efficient clinical and translational research:

- 1) **Research workforce**, or the lack of qualified investigators, mentorship, and sufficient academic reward systems
- 2) **Research operations**, including the high costs associated with research, lack of research funding, regulatory burden, fragmented infrastructure, lack of willing research participants, and incomplete information technology infrastructure

- 3) **Organizational silos**, including a lack of communication and coordination between basic and clinical investigators, lack of interdisciplinary centers within institutions, and differing departmental policies and procedures (16).

The review goes on to highlight Clinical and Translational Science Award awardees who have addressed these issues. Pertinent to the focus of this dissertation, item 2: research operations, is the adoption of centralization of clinical trial operations support (*ibid*). Work through the Clinical and Translational Science Award mechanism has provided numerous evaluations regarding clinical trial metrics over the past several years. The proceedings from a Clinical and Translational Science Award Clinical Research Management Workshop in 2010 listed several metrics that can be used to evaluate where delays in the clinical research process are occurring (16). This report also highlighted that academic medical centers, including those with a Clinical and Translational Science Award, may have prolonged rather than streamlined processes in both study activation and participant enrollment (*ibid*).

In addition to the Clinical and Translational Science Award, the study of cancer is nationally guided. The NCI is the largest branch of the NIH and hosts a funding mechanism for centers devoted to the highest levels of basic and clinical cancer research. This P30 mechanism is called the Cancer Center Support Grant, or CCSG. Centers who are granted a P30 grant by the NCI are deemed NCI-designated cancer centers and those that meet the stringent NCI guidelines in basic, translational, and clinical cancer research are designated as NCI comprehensive cancer centers.

1.4 Gaps in Literature

Despite the focus of the Clinical and Translational Science Award Consortium on issues surrounding clinical research management, there remains a gap in knowledge regarding what constitutes the proper workload, or staff-to-protocol ratio, which most efficiently activates a clinical trial. Likewise, the Cancer Center Support Grant has guidelines for tracking clinical research and centralizing the reporting of cancer research within designated centers, but, like the Clinical and Translational Science Award, there is no framework provided for the determination of feasible clinical research at the site level or guidelines for the efficient activation of clinical trials.

Given the number of studies that indicate many trials accrue poorly at the site level, and the lack of literature regarding quantitative and objective measures to calculate clinical trial accrual *a priori*, it appears that there is a gap in knowledge on **how** to objectively determine which studies to activate. A comprehensive search of the literature, described in greater detail in Chapter Two, found that very few studies focused on prospective, predictive modeling of clinical trial accrual at the site level. The majority of literature was reactive or, at best, looking at patterns prior to the end of the recruitment period but after study initiation. The current research focuses on modifying the clinical trial activation process at the national level (with sponsors). As a single clinical trial site is unlikely to impact the trial activation or accrual status at a national level (the site is one small, replaceable cog in the large, intricate clinical trial machine with multiple parts), more attention has been paid to suggesting change that will have more sweeping impact.

The need for such tools and adaptation to the ever-evolving clinical research landscape was highlighted in another IOM report focused on clinical trial economics. The authors summarized that “[w]hile the technology for gathering clinical data for research has evolved over time, the business model supporting this technology clearly has not evolved with each stage of technology transformation” (40). The report goes on and calls for the development of a new business plan/framework, which was earlier suggested by Denny, who stated that “fatal flaws” would be avoided if clinical research studies were run more like a business (22).

1.5 Previous Work

At the University of Arizona Cancer Center (UACC), we sought to create and validate a predictive model with the outcome of anticipated accrual to be used when considering a prospective clinical trial. To do this we conducted a retrospective cohort study using registry data from the OnCore® clinical trial management system (Forte Research Systems, Inc., Madison, WI) and ClinicalTrials.gov for treatment and supportive care interventional trials.

In univariate analysis, use of an investigational drug, disease team, number of national sites, use of local IRB, number of total months open nationally, months of accrual already completed, and overall proposed national enrollment were significantly associated with accrual. In multivariate analysis, disease team, proposed national enrollment, number of sites, use of local Institutional Review Board (IRB), number of total months open, and number of months already opened were significantly and independently associated with accrual. The full model was significant ($p < 0.001$) and predicted accrual at 94% of actual, maintaining predictive value at multiple cutoff

values. We then validated the model, which correctly predicted whether a study would accrue at least four subjects 75% of the time with a sensitivity of 70.0% and specificity of 78.1%.

The validation of this model shows it to be an accurate, quick, and valuable metric in assessing trial success as well as planning resource allocation and clinical trial costs. This model was incorporated into a feasibility review process at the University of Arizona Cancer Center as one factor in determining whether a prospective clinical trial should be activated locally. This work was presented as a poster at both the 2014 and 2015 American Society of Clinical Oncology (ASCO) annual meetings and a peer-review publication is forthcoming.

1.6 Proposed Frameworks

Industrial-organization (I-O) psychology frameworks exist that have principles which apply to clinical research. Griffith and Sawyer discuss the positive impact that centralized multi-disciplinary teams (such as the kind that are observed in clinical trial offices) have in project success (41). Bertolotti et al. describe the relationship between Multiple Team Membership (MTM), an I-O psychology framework, and staff productivity as an inverse curvilinear relationship (42). Their work utilizing scientific research and development (R&D) teams showed there is an optimal point that balances multitasking by a team member and maximum output. Schultz et al. also demonstrated positive impact in the performance in surgeons within the R&D space when multitasking was utilized (43). Applying this to clinical research, where staff often juggle multiple clinical research studies, this framework can be utilized to examine what is the optimal workload for staff working on study activation. It is important to identify not

only the optimal quantity of projects, as highlighted by Bertolotti, but also the organizational context of the types of projects overseen by the team member (44). As there are no national standards for clinical trial office staffing, several different models exist, particularly around regulatory management and study activation. Some institutions utilize a dedicated study activation team while others utilize the same team to perform study activation and regulatory maintenance. Still others utilize a “cradle-to-grave” approach, where clinical research coordinators perform both regulatory and subject management tasks. It is **unknown** whether staff devoted to regulatory activities are more efficient at opening studies that will ultimately accrue participants than team structures that have a more diverse portfolio of study management. Thus whether team members who work on both regulatory and study coordination are more or less productive than teams who focus solely on regulatory is an aim of this dissertation project.

In addition to multitasking and team management, project management framework regarding infrastructure analysis can also be applied to clinical research. Stage-gate modeling describes the need to put parameters in place to “kill,” or abandon development of, a project that is not going to succeed (45). Entities such as the U.S. Department of Health and Human Services (DHHS) utilize this framework to ensure that entity needs are met throughout the development of a project (46). The mechanism for this model includes formal processes that utilize experienced staff members to make decisions on whether a project should proceed to the next step. This is done multiple times (standards set three or five stages, each separated by a decision point, or gate) during project development, allowing for projects that do not appear to be hitting preset benchmarks to be objectively removed from development, allowing projects with maximum

potential for success to be fully developed. The stage-gate process is particularly useful when there are too many projects that need to be completed quickly, such as the simultaneous activation of multiple clinical trials. Under this framework, formal and objective decision points are preferred as the entity is guided on projects to undertake from their customers. In the clinical trial arena, the entity would be the clinical trial office that is charged with activating the study; the customer is the investigator requesting the project to be activated. Stage-gate processes are useful because, without them, there is no mechanism to stop a project from initiating. Applying this to the clinical trial activation process, many institutions do not have a mechanism to halt or abandon activation of a project. Once an investigator (the customer) requests a project to be activated, there are no established and systematic criteria to determine that it is not feasible to continue activation at the site level. The exception to this is in the NCI Cancer Center Support Grant (CCSG) guidelines for NCI-designated cancer centers. The Protocol Review and Monitoring System (PRMS) is given authority to reject a project based on scientific merit and/or feasibility. However, no national guidelines or tools exist for systematic and objective decision-making by this body. Another gate exists in the clinical research study activation process. The Institutional Review Board (IRB) is required to review human subjects research (including clinical trials) prior to study activation and has the authority to deny studies from opening at a center. However, this body approves studies based on factors related to ethics and subject safety and welfare, not logistics or feasibility unless it impacts the rights, safety, or welfare of research subjects.

The stage-gate framework, partnered with the Multiple Team Membership (MTM) framework, is highlighted by Griffith and Sawyer in their work discussing productive R&D teams (41). The R&D team process is similar to clinical research teams, in that they are both processes with outcomes that are scientifically motivated but have financial considerations in both monetary and human resources. Additionally, staff and leadership in both areas are often scientifically trained, not business trained, and thus may not include project management strategies into decision-making workflows. Therefore the combination of these two frameworks appears to be an optimal methodology to assess and streamline processes within clinical trial activation.

1.7 Policy Implications

This work addresses concerns raised by the IOM and independent researchers regarding the lag in the learning healthcare system, or applying research to clinical practice. By streamlining activation processes at the site level and activating studies that will be successful in accruing subjects, the clinical development process is expedited, reducing research waste.

1.8 Specific Aims

Study Population: The scope of this dissertation work is cancer centers adhering to the NCI Cancer Center Support Grant (CCSG) guidelines. This population provides a homogenous population of study sites that conduct a large volume of clinical research with support staff dedicated to conducting clinical research (and not involved in clinical practice). These centers all have centralized clinical trial offices that manage some aspect of clinical research under guidelines outlined by the NCI, providing a base level of standardization. The use of this

population will allow conclusions to be drawn regarding the impact of processes and tools on workflow and defined outcomes (accrual and study activation) without the potential confounding factors of clinical practice. Additionally, as the pool of cancer protocols is large—stated earlier as approximately one-third of all currently-open clinical trials on ClinicalTrials.gov—a more homogenous pool of clinical research protocols can be assessed, reducing the potential for bias due to differential selection (e.g. research in different disease states impacts the outcomes of interest differently due to the nature of their design).

1.8.1 Specific Aim 1

Aim: Utilizing stage-gate framework, assess the impact of a feasibility committee on the study activation workflow for the clinical trial enterprise at the site level.

- Hypothesis 1a: There is no difference in total site accrual per study between cancer centers that have formal and specific processes to make “Go/Abandon” decisions (a feasibility committee, followed by PRMS, and IRB), fitting the 5-stage/3-gate framework theory, versus those that do not, using the 3-stage/2-gate process (PRMS and IRB review only).
- Hypothesis 1b: There is no difference in the number of studies that have zero accrual between cancer centers that have formal and specific processes to make “Go/Abandon” decisions (a feasibility committee, followed by PRMS, and IRB), fitting the 5-stage/3-gate framework theory, versus those that do not, using the 3-stage/2-gate process (PRMS and IRB review only).

Rationale for Aim: The current clinical trial literature repeatedly calls for innovation in processes surrounding study activation and the development of research business plans. Frameworks for such plans have been established in the industrial-organization (I-O) psychology literature studying the R&D industry. These studies report that the utilization of formal decision-making processes increases the ability to move profitable projects forward and reduce the investment in projects that ultimately fail. Some cancer centers have independently developed feasibility committees, which make these “Go/Abandon” decisions regarding study activation. Since all cancer centers must perform PRMS and IRB review, assessment of what impact a feasibility committee makes on activation timelines and study success (ability to accrue) can be measured.

Summary Methodology: Utilizing administrative and protocol data collected from participating cancer centers, assess whether centers adhering to more formal review processes (a 5-step framework, incorporating a feasibility committee) have higher accrual per protocol and fewer zero accruing trials.

1.8.2 Specific Aim 2

Aim: Utilizing Multiple Team Membership (MTM) framework, investigate the relationship between study activation times, number of staff, staff full-time equivalent (FTE), protocol accrual, and the protocol workload, defining workload as clinical trial protocols to be opened and defining activation time as an indicator of success (fewer number of days from start to activation = more successful activation).

- Hypothesis 2a: There is no relationship between time to study activation and protocol workload (number of protocols per staff member).
- Hypothesis 2b: There is no relationship between overall protocol accrual and protocol workload (number of protocols per staff member).
- Hypothesis 2c: There is no relationship between the percentage of zero-accruing protocols and protocol workload (number of protocols per staff member).
- Hypothesis 2d: There is no relationship between time to study activation and percent full-time effort (% FTE) dedicated to study activation.
- Hypothesis 2e: There is no relationship between overall protocol accrual and percent full-time effort (% FTE) dedicated to study activation.
- Hypothesis 2f: There is no relationship between the percentage of zero-accruing protocols and percent full-time effort (% FTE) dedicated to study activation.

Rationale for Aim: Multiple Team Membership (MTM) framework suggests an inverse curvilinear relationship between the number of projects a team member is a part of and efficiency. This framework has been applied in the R&D industry, measuring the efficiency of specialized teams in science. As the R&D industry is similar to the clinical research enterprise, in that they are both science focused, technologically advanced, and require specialized team members, this framework should be applicable to the clinical trial team structure seen within cancer centers.

Summary Methodology: Utilizing the MTM framework, assess study activation workload two ways. The first will be by the number of staff members participating in regulatory activities. The second will be by the number of FTEs working on regulatory functions. Utilizing both measures will inform on two workload perspectives: 1) determining what staffing model is most efficient (dedicating staff to regulatory functions or having the same person manage all aspects of the study), and 2) what level of workload maximizes the outcome of interest. Protocol workload (number of protocols per staff member) will be plotted against the following outcomes: 1) overall accrual. 2) percent zero-accruing protocols, and 3) time to activation to determine at what level of workload maximizes the success of a protocol (clinical trial accrual) while minimizing activation time. Multiple regression models will be utilized to control for factors that may confound the relationship between the dependent and independent variables of interest for each specific hypothesis.

1.8.3 Specific Aim 3

Aim: Determine whether a statistical model utilizing protocol characteristics known prior to study activation can be created to accurately quantitate clinical trial accrual for participating sites.

- Sub-Aim 3a: Using independent variables associated with clinical trial accrual, create one overall model to predict clinical trial accrual.
- Sub-Aim 3b: For each site, assess whether the site-specific model predicts clinical trial accrual better than the overall model.

- Sub-Aim 3c: For each site, assess whether the model-predicted accrual is more accurate than disease team-predicted accrual.

Rationale for Aim: It is not in the best interests of the center (financial or scientific) to open a protocol that will not accrue subjects to trial. A model utilizing past performance data gives an objective, quantitative measure to determine whether a study should be activated. A regression model fits the framework presented by Barnard for the use of predictive models in that it would be:

- 1) Simple to use and understand (the outcome is a single number with actual meaning in the clinical trial enterprise)
- 2) Able to adapt to epidemiologic changes (through the modification of independent variables)
- 3) Able to adapt to environmental changes (through the modification of independent variables)
- 4) Able to take account of center recruitment
- 5) Able to inform commissioning decisions (47).

This model, incorporating variables that affect clinical trial accrual that are known *prior* to opening a trial, should accurately assist in objective decision-making prior to investing significant resources in study activation.

Summary Methodology: For individual institutions as well as to complete Sub-Aim 3a, univariate analyses will be performed to determine which protocol- and site-specific

characteristics are significantly associated with clinical trial accrual. Reduced models of related independent variables will be performed to test for effect modification. Statistically significant variables will be included in a full model. Actual accrual will be correlated with predicted accrual. A mixed model using institution as the random effect will be used to account for the hierarchical nature of the data. Institution is picked as the random effect as each institution will have infrastructure that independently (from other centers) and consistently (within the institution) affects the conduct of clinical trials. Measures of fit, such as pseudo- R^2 , AIC, and BIC values will be utilized to assess which model best fits the data.

To complete Sub-Aims 3b and 3c, both overall model (unadjusted and adjusted) predicted accrual and disease team-predicted accrual will be compared to the site-specific models and actual accrual using Pearson's correlation coefficient, percent matching categories, sensitivity, specificity, and overall accuracy to assess which method is more accurate.

CHAPTER 2. LITERATURE REVIEW

2.1 Introduction to the Issue and Roadmap

Clinical research is the backbone for medical practice, as highlighted by the following statement from the Institute of Medicine (IOM) in 2010: “[e]fficiently generating medical evidence and translating it into practice implies a ‘learning health care system’ in which the divide between clinical practice and research is diminished and ultimately eliminated” (1). However, the process of obtaining evidence to apply to clinical practice is slow and riddled with bottlenecks (further explored in section **2.2 Clinical Research Policy Issues**). The clinical trial process accounts for 97 of the 128 months it takes for a single agent to traverse the process to gain FDA approval (5). During this time, an agent goes through at least three clinical trials (phase I, II, and III) which examine the agent’s safety, efficacy, and effectiveness. These studies involve hundreds or thousands of people and are conducted at centers across the U.S. and the globe. The IOM estimates that approximately 90% of phase III clinical trials, the largest studies which are pivotal for FDA approval, do not meet their planned timelines for recruiting human subjects (1), elongating the time it takes to obtain data and often requiring more centers to conduct the study than originally planned. This increases the cost of drug development (section **2.3 Clinical Trial Economics and Workload**). Each center participating in a clinical research study must go through a lengthy activation process to open a study, including institutional, sponsor, and federal requirements, such as budget and contract negotiation and IRB review (section **2.4 Study Activation**).

Despite center investments in these studies, a large proportion of studies (reported between 20 and 50%) do not enroll a single subject at the center level (1, 18, 25-26, 48; section **2.6 Clinical Trial Accrual**). This equates to a significant investment of human and monetary resources to open these clinical trials with no financial or scientific return. Four main categories of issues are associated with clinical trial accrual (section **2.6.1 Factors Affecting Clinical Trial Accrual**), three of which can be addressed, to varying extents, prior to clinical trial activation. These are protocol issues (section **2.6.1.1 Protocol Issues**), site issues (section **2.6.1.2 Site Issues**), investigator issues (section **2.6.1.3 Investigator Issues**), and participant issues (**Figure 1-1**). However, these issues affect accrual differently based on the stage of a clinical trial. While changes at any level will impact accrual (both positively and negatively) throughout the life of a study, changes to different categories impact accrual greater at different stages of study development, activation, and execution (**Figure 2-1**). It is crucial to make choices that maximize accrual early on in the activation process to make the greatest impact. Thus focus should not be on factors that affect accrual after a study has been activated at a center, mainly participant and investigator issues, but on maximizing potential accrual prior to a study opening through understanding and choosing studies that are most likely to succeed. To maximize resources and accrual to clinical trials, studies that will ultimately not accrue must never be opened at a site. They must be shut down, or abandoned, in the activation process (section **2.4.1 Stage-Gate Framework**). It is the position of this investigator that focus should not be on trying to maximize accrual on studies that are already open to accrual as, despite best efforts, a certain proportion of studies will never accrue subjects because the site is not amenable to recruiting for that particular type of study. If we focus earlier in the process, we can “prevent” the problem of

having unproductive studies rather than trying to “diagnose” and “cure” the study when it is open and not accruing. Thus it is important to know what will succeed before activating a study.

Focus must be on identifying studies that are most likely to succeed at a site based on past protocol and site characteristics of like trials (section **2.6.2 Modeling Clinical Trial Accrual**).

Additionally, human resources, specifically regulatory staff, must be optimized with the proper structure and workload to ensure that studies are activated in the most efficient manner (section **2.5.1.1 Multiple Team Membership (MTM)**).

Clinical research is rapidly evolving with new forms of identifying and treating diseases, such as genomic mapping and targeted therapy; however, clinical trial administration processes have not progressed at the same rate (section **2.2.1 Call for Reform**). The need for adaptation to the ever-evolving clinical research landscape was highlighted in an IOM report focused on clinical trial economics. The authors summarized that “[w]hile the technology for gathering clinical data for research has evolved over time, the business model supporting this technology clearly has not evolved with each stage of technology transformation” (40). The report goes on and calls for the development of a new business plan/framework. In a separate study, Denny suggested that clinical research studies need to run more like a business by making better choices and educating investigators and patients to avoid “fatal flaws” (22). His analysis stated that a business plan would improve recruitment rates, allow studies to finish faster, allow the proposed scientific questions to be answered, achieve the statistical goals of the study, improve patient care due to results being published more quickly and new studies designed with updated goals, as well as improve study costs and/or efficiency rates by freeing up needed and expensive resources (*ibid*).

The purpose of this dissertation project is to embrace the creation of a site-level business plan for the activation of clinical trials through defining an infrastructure that 1) most efficiently uses staff resources to activate studies (section **2.5.1.1 Multiple Team Membership (MTM)**) and 2) only opens studies that are going to succeed (sections **2.4.1 Stage-Gate Framework** and **2.6.2 Modeling Clinical Trial Accrual**). This project aims to study objective measures to fulfill both of these goals through the analysis of the past performance of protocols conducted at cancer centers as well as their study activation processes and staff resources to map out study types, staffing structures, and workloads that maximize study activation efficiency (reducing the time it takes to open a study, thus maximizing the time available to accrue subjects) as well as choosing studies that are most likely to accrue subjects. To do this, existing methodologies within the biostatistical and industrial-organization (I-O) management will be applied to clinical research and their ability to measure outcomes assessed. Then, when the “point of no return” in study activation is reached—opening a study to accrual—and the investment of extensive staff resources is made, there can be confidence that the most informed decision has been made. Then, a center will positively contribute to the national completion of a clinical trial while minimizing research waste.

2.2 Clinical Research Policy Issues

Many of the agents studied for clinical effectiveness fail to make it to the stage of FDA approval and public dissemination. The Tufts Center for the Study of Drug Development (CSDD) reported in November 2014 that 80.3% of compounds going through the drug development process are discontinued at some point during the almost 11-year process of study and approval.

At the time of the report, only 7.1% of the 1,442 compounds reviewed had received FDA approval (5). A contract report to the U.S. Department of Health and Human Services (DHHS) summarizes many of the common barriers to conducting clinical research. These include high monetary costs, length of time to conduct the trial, recruitment and retention difficulties, clinical research workforce insufficiencies, sponsor-imposed barriers, and regulatory/administrative barriers (10). In early 2014, the Lancet published a series of articles focused on research waste (3, 6, 27, 49-51). They reported that it takes over 20 years to translate basic science into clinical practice with very few discoveries widely utilized in the clinical realm (6). Another report quantitated the percent of original research that made it into clinical practice as 14% with an average lag of 17 years before reaching clinical practice (52) with the U.S. expenditure on life science research in 2010 at \$240 billion, just over 5% of the U.S. health care expenditure (49). Chalmers et al. estimated in 2009 that 85% of research expenditures were classified as waste and categorized research waste into four categories:

1. Choosing the wrong research question
2. Doing studies that are unnecessary or poorly designed
3. Failure to publish relevant research promptly, or at all
4. Biased or unusable reports of research (53).

Discussed in the commentary by Macleod et al., the biomedical research process is highlighted by interactions that are made within complex, independent systems that act interdependently (49). The first category of research waste discusses the implementation of research that does not address a question of interest to patients; otherwise, a lack of a patient-centric approach to

research development. The same focus can be applied to site-level selection of clinical trials and applicable to this dissertation. Choosing the wrong question or, in the site's case, the wrong clinical trial for activation, means that resources are placed in a clinical trial that will not accrue adequately at the site. Those resources are drawn away from another trial that could impact patients at the site more directly, ultimately impacting accrual due to an inability to find eligible subjects or to gain interest by eligible patients to the trial.

Focusing back towards the 2014 Lancet series, Chalmers focuses on the research waste that occurs when decision makers (e.g. policymakers, patients, and physicians) are ignored in the research development process (6). Another source of waste, highlighted by Ioannidis et al. in the same series, was the inability to obtain quality information from research due to a lack of statistical precision or power (27). A lack of statistical precision or power can be attributed to low accrual. Poor statistical precision or power can lead to an inability to reproduce research results and an inability to inform clinical practice in an intelligible way using quality research evidence. This is highlighted by the findings that the pharmaceutical company Bayer could not replicate 64% of 67 academic oncology and cardiovascular studies reported in peer-reviewed publications (54). Similar statistics were reported by Amgen, as they reported 89% of oncologic pre-clinical findings for potential drug targets relevant to clinical practice could not be verified (55).

The third article in the Lancet series focused on waste due to the regulation and management of research. Al-Shahi Salmon et al. identified several areas that lead to research waste, including

trials being overlooked or unnecessarily run due to research being underpowered, research performed too slowly, or research being too costly (3). One of the areas highlighted by this group was that the regulatory process is “burdensome and too slow,” especially when delays are due to consecutive approval processes instead of concurrent processes. Numerous reviews lead to replication of similar processes, information giving, and additional time by researchers to address a medically relevant question. Additional issues that lead to inefficiency, and thus research waste, is slow recruitment and retention of research subjects. An extended recruitment period leads to a delay in getting valuable information regarding treatment options to the public and increasing the amount of funding needed to complete the study. The review states that “many processes... intended to improve the quality and safety of clinical research are costly, time-consuming, and of unproven effectiveness (3).”

Al-Shahi Salmon et al. suggested four recommendations for reducing research waste regarding regulation and management. Two recommendations are pointed towards research regulators and include recommendations to reduce other causes of research waste and inefficiencies as part of their regulating and monitoring processes, particularly surrounding study design, and to harmonize laws, regulations, and guidance surrounding research conduct. The last two recommendations are geared towards researchers, research managers, and clinical practice administrators. The two recommendations to increase research efficiency are through using effective methodologies for all aspects of clinical trial practices (e.g. design, recruitment, retention, data collection), to do research on clinical trial efficiencies, and to better improve patient population knowledge of clinical research (3).

Primary audiences for high-quality evidence, as the type found through clinical research, include patients, physicians, payers, purchasers, health care administrators, and public health policymakers. Research waste lowers the quality of clinical research. In 2003, Tunis et al. discussed the impact on health policy due to the lack of quality, published clinical research (56). The clinical practice gaps that remained unaddressed, such as variance in clinical practices and high rates of inappropriate care, led to their conclusion that systematic flaws exist in the clinical research program that affect the ability for policymakers to make decisions. The prevalence of knowledge gaps, especially gaps formed by the lack of studies examining the efficacy of newer treatments against established treatments, leads to recommendations for health care and treatment that are not definitive. This results in inefficiencies in the health care system, including delayed coverage of technologies by insurance companies and inefficiency of recovering the cost of research and development of new technologies and/or products (56).

The cost of research, especially large studies, is prohibitive, especially to federal funders and institutions. Thus, when such extensive funding is granted, high-quality evidence is needed to justify the monies (often tax-payer monies) spent. This means, at the protocol level, that clinical trials must accrue the needed number of subjects to be powered appropriately to meet their endpoints and be executed in a timely fashion to ensure that findings are still relevant.

2.2.1 A Call for Reform

The IOM stated in a report regarding the NCI cooperative group program that “the current structure and processes of the entire clinical trials system needs to be redesigned to improve

value by reducing redundancy and improving the effectiveness and efficiency of trials (2).” In addition to the IOM’s call to the NCI to reform their cooperative group program, in 2012 the NIH asked the IOM to evaluate the Clinical and Translational Science Award (CTSA) program, an infrastructure grant that supports clinical research (among other things), and provide recommendations for improvement. While not a direct clinical trial sponsor, the Clinical and Translational Science Award (CTSA) mechanism provides external infrastructure support to clinical trial sites. The 2013 IOM report on these awards contains the findings of this evaluation (57). Within the numerous recommendations that the IOM makes regarding the program, two are relevant to this project:

1. “Build on the strengths of individual CTSA’s across the spectrum of clinical and translational research” in that the CTSA’s should “drive innovation and collaboration in methodologies, processes, tools, and resources across the spectrum of clinical and translational research”
2. “Formalize and standardize evaluation processes for individual CTSA’s and the CTSA program using clear, consistent, and innovative metrics... that go beyond standard academic benchmark of publications and number of grant awards.”

These views were echoed outside the IOM by Chalmers et al. who, in the Lancet series on research waste, stated that prioritization of research resources through communication, analysis of health systems, research funding, and research feasibility are needed to reduce waste (6). In the same series, Al-Shahi Salmon et al. recommended that “[r]esearchers and research managers should increase the efficiency... in research through the use of research designs known to reduce

inefficiencies, and do additional research to learn how efficiency can be increased.” One example of increasing efficiency was provided in the IOM report regarding Clinical and Translational Science Awards (CTSAs) and applauded improvements to the clinical research enterprise through changes in institutional review board (IRB) processes such as single IRB review for collaborative research projects (57).

2.3 Clinical Trial Economics and Workload

In addition to the low number of compounds achieving FDA approval, the costs associated with drug development and testing are growing at an approximate 9% annual rate. The IOM reported that a 14,000 subject/300 research site global trial can cost \$300-600M (U.S., 2009; 1). Clinical trials have fixed and variable costs that must be covered by the site conducting the trial. Fixed costs to conduct clinical trials fall into one of two categories: 1) non-protocol fixed costs (e.g. staff, administration, electronic equipment) and 2) protocol-specific fixed costs (e.g. study activation costs). Both types of costs are incurred regardless of the number of subjects ultimately accrued and, in a fully functioning system, are recuperated through successfully accruing the contracted number of research subjects to clinical trial, for industry-sponsored clinical trials. Oftentimes, the majority of payment for industry-sponsored clinical trials is based on subject accrual, negotiated prior to the start of the study, and must cover the site's fixed and variable costs for the entire life of the study. For studies that do not accrue any research subjects, minimal payments are received and usually only partially cover the costs associated with activating the study locally, not staff time to maintain a protocol.

Funding for government-sponsored clinical research differs from the industry model. Grant-funded clinical research has a budget proposed by the site that accompanies the grant application. This budget is then adjusted and approved by the federal funder. At the site level, when adjusted for inflation, federal funding for sites is decreasing (58). The IOM reports it is estimated that the NIH funds centers 20-40% less than what it costs to actually conduct a clinical trial (1). Many federally funded cancer clinical trials are sponsored through the National Clinical Trial Network (NCTN), an NCI-hosted clinical trial network, the 2013 replacement of the cooperative group system. This network has set per-subject reimbursements (averaging \$2,000-\$4,500 per subject), which are significantly less than the cost to conduct a trial (Tate, unpublished). This decrease in both funding and investigators makes it important for sites to choose studies that contribute positively to the scientific mission or economic status of the center.

The cost of clinical trials has been an issue for over 25 years. Piantadosi et al. published one of the first clinical trial models in 1987, developed to estimate the cost of a randomized, controlled trial (59). Buonasegna reported that economic reasons were the reason almost 10% of clinical trials failed. For example, an expected cost of development greater than projected profits (from the sponsor's perspective) could lead to abandonment of a drug, even if it appears to be efficacious (31). In 2003, Emanuel et al. reported a simulation study using 21 institutions (including two academic medical comprehensive cancer centers) that modeled the time and associated costs with conducting clinical research. Their findings showed that, on average, it took about 4,000 hours (range 1,512-15,699) of work to conduct government-sponsored and industry-sponsored trials that accrued 20 subjects and included 17 study visits, equating to

approximately 200 hours of effort and \$6,100 per subject (U.S. dollars, range \$2,098-\$19,285; 1999; 60). Specifically, the time to activate a study was estimated to be 57 hours for NIH-sponsored research and 117 hours for industry-sponsored research (*ibid*). A National Bureau of Economic Research (NBER) grant examined the site work effort (SWE), a patent-pending calculation of protocol complexity, which showed that between 1989 and 2011 workload increased approximately three-fold. Oncology made up 9% of the protocol contracts in this database (5.89% in 1989, increasing to 12.55% in 2011; 17). The IOM states in their 2010 report the findings of two studies that estimated median costs of \$5,000 (60) to \$6,000 (35) per subject, despite the reimbursement rate of \$2,000 for cooperative group protocols, a figure which, until 2014, had remained constant since 1999 (2).

Another study found that the average research cost per subject (from the sponsor's perspective) on a phase III oncology clinical trial is approximately \$75,000-125,000 (U.S. 2013). This is up from \$3,000-\$5,000 in the early 1990s and, because of individual site issues, it is not uncommon to see more trial sites than planned patients to enroll (61). However, this cost is not paid out to the site. Sponsor payments to institutions are negotiated and contracted, therefore information regarding average reimbursement/site payment levels are not available. The NBER report extensively explored the cost of industry clinical trials, using studies reported in the MediData system between 1989 and 2011 to calculate the total grant cost per subject to sites. During this period the authors found that clinical trials had an average annual growth rate of 7.5% in the total grant cost per subject (calculated as all payments made by a trial sponsor to a site that enrolls research subjects), increasing from \$3,773 to \$16,567, or approximately four-fold, during the

reviewed time period (62). The median total grant cost per subject was lower, ending at \$13,222 in 2011 (*ibid*).

2.4 Study Activation

In order to conduct a study at a site, it must go through a process of approvals and procedures to ensure the site is prepared to conduct the study. As many phase III studies require dozens if not hundreds of sites, the site-level activation time contributes to a notable portion of the time devoted to the clinical testing phase of drug development. In 2010, Dilts et al. reported that a phase III cooperative group protocol required 769 steps, 36 approvals, and a median of 2.5 years to go from formal concept review to study opening (18). Ninety-five (95) of these steps were at the comprehensive cancer center level (*ibid*) and included time-consuming processes such as scientific review at the institution level (Protocol Review and Monitoring System; PRMS), IRB approval, and budget negotiations. Another study of the Cancer and Leukemia Group B (CALGB) national group mapped 370 distinct processes with a median time from initial concept to study activation of 784 days (range: 537-1130 days; 19). With this information it is not surprising that only about 60% of cooperative group trials are completed and published (2). When mapped it took, on median, 171 days (95% confidence interval: 158-192 days; range: 27-657 days) to open a single clinical trial at one site and 50% of the steps/processes towards opening an oncology clinical trial held no value (20). These institutional activation timelines contribute to the overall drug development time, which is 7.8 years for the phase I-III clinical investigations for a single agent (15). It is also reported that the length of this overall process has increased over time (16).

Emanuel et al. reported a simulation study using 21 institutions (including two academic medical comprehensive cancer centers) that modeled the time and associated costs with conducting clinical research. Their findings showed that, on average, the staff time needed at a single site to activate a study was estimated to be 57 hours for NIH-sponsored research and 117 hours for industry-sponsored research (60). Another study reports that it takes 49 days longer to initiate a clinical trial at an academic center than at professional or community sites (16). An analysis of over 24,000 protocols showed that workload increased approximately three-fold between 1989 and 2011, with a mean annual increase of 5.2% while, during the same time period, the number of patients per site has decreased from a median of 20 to a median of eight (17). This equates to a significant time and salary investment prior to any subjects being enrolled on a protocol.

In 2011 the IOM held a two-day workshop focused on the foundation of the clinical trial enterprise, the randomized controlled clinical trial (RCT) due to a “widening separation” between clinical research and clinical practice (21). This multidisciplinary panel of clinical research experts representing academia, the federal government, non-profit organizations, and pharmaceutical companies, called for the development of research business plans to speed cultural changes in clinical trial organization. A main finding was that, while research technologies (e.g. genetic mutation screening, cancer sub-typing, and updated study designs) have advanced at a rapid pace, research administration lagged needed cultural changes to administer these more advanced clinical research trials (*ibid*). This view was corroborated in a separate report stating that increased regulatory processes, inconsistencies within these

processes, and the timelines to conduct both basic and clinical (observational and interventional) research creates waste (monetary, human workforce, and time resources; 3).

Steensma publically called out the NCI as impeding the progress of clinical practice through the impeding of clinical research with bureaucracy, specifically imposing many unnecessary regulatory requirements which extend the time period to conduct clinical research (61). In response the NCI discussed the improvements to the clinical trial development and implementation process through the 2008 Operational Efficiency Working Group (OEWG). The constitution of this group included representation from the cooperative groups, NCI-designated cancer centers, patient advocates, community oncologists, industry, and the FDA. In their two-year effort their mandate was to create a comprehensive strategy to increase efficiency in clinical trial development (63). One main recommendation was to no longer pursue trials that took longer than 18 months (for phase I/II trials) or 24 months (for phase III trials) to activate (64). In addition, project managers were instated to oversee the activation process and timelines for the trial at a national level. The working group found that, by implementing the new deadlines, as well as a joint NCI-investigator conference call after review of the concept and again after protocol development, development time decreased by 18.3%, to a median of 442 days, for phase I/II studies and 45%, to a median of 395 days, for phase III studies. Despite implementing the more stringent activation timelines, only two of the 159 phase I/II trials and none of the 25 phase III trials were terminated during the development stage for exceeding the absolute deadlines (64). However, there are not yet any reports discussing the impact, either positive or negative, of this change, on other metrics such as site activation timelines or overall trial accrual.

An IOM review of the Clinical and Translational Science Award (CTSA) program categorized three site-level barriers to efficient clinical and translational research:

- 1) Research workforce, or the lack of qualified investigators, mentorship, and sufficient academic reward systems
- 2) Research operations, including the high costs associated with research, lack of research funding, regulatory burden, fragmented infrastructure, lack of willing research participants, and incomplete information technology infrastructure
- 3) Organizational silos, including a lack of communication and coordination between basic and clinical investigators, lack of interdisciplinary centers within institutions, and differing departmental policies and procedures (16).

This report also highlighted that academic medical centers, including those with a Clinical and Translational Science Award (CTSA), may have prolonged rather than streamlined processes in both study activation and participant enrollment. One method that Clinical and Translational Science Award (CTSA) awardees have used to address these issues, particularly issues related to research operations, is the centralization of staff support for clinical trial operations; however, there has yet to be developed a metric of workload for these centralized offices to optimize study activation timelines (16).

2.4.1 Stage-Gate Framework

This project seeks to provide a process for study selection and workload management during study activation. Established frameworks regarding project feasibility within the industrial-

organization (I-O) psychology literature have been validated in the Research and Development (R&D) industry. Specifically, the stage-gate framework addresses Denny's suggestion that clinical research needs to run more like a business in that better decisions need to be made regarding which clinical trials to activate (22).

Stage-gate modeling is suggested for use when there are many competing projects vying for limited resources in a constrained time period. The model provides two goals by which its processes drive new project resource allocation. First, it allows the allocation of needed resources to projects that are likely to succeed. The second goal of stage-gate modeling is to put parameters (gates) in place to abandon development of a project that is not going to succeed (45). Entities such as the U.S. Department of Health and Human Services (DHHS) utilize this framework to ensure that entity needs are met throughout the development of a project (46). This model includes formal processes (gates) that utilize experienced staff members to make decisions on whether a project should proceed to the next step ("Go/Abandon" decisions). This is done multiple times during project development, allowing for projects that are not hitting preset benchmarks to be objectively removed from development, allowing projects with maximum potential for success to be fully developed. Applying this to the clinical trial activation process, many institutions do not have a mechanism to halt or abandon study activation. Once an investigator (the customer) requests a project to be activated, there are no established and systematic criteria to determine that it is not logistically feasible to continue activation at the site level.

This first gate is a feasibility committee. In the clinical research space, this equates to the pitch of a proposed clinical trial. The protocol would be brought to an administrative committee that would review the trial to determine whether the institutional resources, whether financial, structural, or human, are available. The benefit of a feasibility committee is that all potential ideas are considered consistently and allow fundamental research to be 1) directed, 2) focused, and 3) productive. Valid ideas are given resources to activate. Ideas without merit or that need further consideration can be set aside in lieu of more viable projects (45). The authors of stage-gate modeling specifically recommend the stage-gate process for research processes with the criteria of passing the gate to be strategically focused. Five criteria are suggested to be reviewed for fit of the project before passing to the next stage:

1. The degree of strategic fit and importance to the entity
2. The ability to achieve strategic leverage (e.g. growth, impact)
3. The potential for reward
4. The likelihood of technical feasibility
5. The likelihood of commercial success.

These criteria fit the mission of the academic medical center and cancer centers in that:

1. The degree of strategic fit and importance to the entity would be how the project meets the mission and vision of the center/institution
2. The ability to achieve strategic leverage (e.g. growth, impact) equates to the site's ability to meet the needs of their community
3. The potential for reward in the way of scientific credit and progressing clinical science

4. The likelihood of technical feasibility through the availability of appropriate human and physical resources to conduct the trial and meet accrual goal(s)
5. The likelihood of commercial success to further the scientific mission of the center through the production of new intellectual property for the clinical treatment of disease or conditions.

Stage-gate modeling partners well with portfolio management (45), a strategy that clinical research teams often utilize at regular intervals. Stage-gate modeling suggests the following four evaluations be done to optimize a portfolio of projects:

1. Selecting high value projects
2. Achieving the right balance of projects
3. Selecting the right number of projects
4. Strategic alignment.

Regular evaluation of the portfolio, even outside of a formal “gate” in the activation process, will ensure that only projects that are highly likely to succeed (accrue an appropriate number of subjects) will continue through the activation process. Activating a project does not end the need for evaluation within the portfolio. Ongoing projects that are not progressing as expected or are not as high of a priority as a new study in the activation pipeline may be closed through these regular evaluations to ensure that the portfolio stays balanced. Conversely, the activation of a new project may be placed on hold because of an ongoing trial (e.g. a trial that competes for the same patient population may not be wise to activate as it will drain the patient population and

neither study will be able to achieve its goals for analysis). However, this is not the focus of this dissertation and not considered further.

2.5 Cancer Clinical Research

Cancer is defined by the NCI as “the name given to a **collection** of related diseases” (emphasis added; 11). Currently cancer is the second leading cause of death in the U.S. and will surpass heart disease by 2030 at current rates (9). In 2013, approximately 584,881 deaths were attributed to cancer (8). While breakthroughs have been made in some forms of cancer, overall incidence and mortality rates in many cancer subtypes are not improving (7). In fact, as molecular testing is becoming more commonplace and genetic mutations identified, cancer is being further differentiated and subtyped. Now, consideration for treatment must not only consider the anatomic origin of the cancer, but the genetic make-up of the tumor.

Over the last 70 years, over \$100 billion (U.S.) have been devoted to the study of cancer by entities such as the NCI and the American Cancer Society. When considering drug development costs of a traditional pipeline consisting of phase I, II, III studies for FDA submission, oncology is the third most expensive area of clinical research. It is estimated that the cost of a singular oncology agent going through this development pathway exceeds \$78 million (U.S.; 10). Currently approximately one-third (32.7%), or about 5,600, of the open, interventional trials listed on ClinicalTrials.gov are studying cancer. Cancer is central to the new Precision Medicine Initiative (PMI), announced by President Obama on January 30, 2015, and administered through the NIH. This initiative seeks to advance patient-driven treatment based on the genetic

uniqueness of the individual's disease or condition and approximately one-third (32.6%) of the Precision Medicine Initiative's budget has been earmarked for cancer with the specific objective of accelerating "the design and testing of effective, tailored treatments for cancer..." (12).

Cancer clinical trials have the highest failure rate (55) and as costs continue to increase and cancer becomes subtyped into more diseases, it is becoming increasingly harder to power clinical trials as populations possessing the correct inclusion criteria shrink due to molecular subtyping. To address this need clinical researchers, including biostatisticians, have developed adaptive designs to assist with accelerating the drug development process while meeting statistical power to address the hypotheses of trials (65). These adaptive trials include central protocols that allow the enrollment of multiple anatomical origins of disease and then sub-divided (either randomly, not randomly, or a combination of both) based on the molecular mutations present in their tumors. Protocols such as the NCI-MATCH trial, an adaptive trial that spans multiple anatomical origins with an estimated 20 arms, proposed enrollment of 1,000 subjects (from 3,000 people screened), and numerous study agents, activated in August 2015, are a national approach to modifying the traditional drug development pipeline process to fit the evolution of clinical science and technology (66).

2.5.1 Clinical Trial Pipeline at an NCI-Designated Cancer Center

Cancer centers exhibiting excellence in providing cancer care and performing research in the basic and clinical spaces are supported through an NCI P30 federal grant called the Cancer Center Support Grant (CCSG). This grant provides significant infrastructure support for

centralized clinical research support offices. These units, called clinical trial offices, or CTOs, provide a core of professional staff dedicated to providing clinical trial support in at least one of the following areas: regulatory support and compliance, clinical research coordination, data management, internal monitoring, research nursing, and/or clinical research informatics.

Many cancer centers have their centralized research support staff within the clinical trial office reporting through two hierarchies. The official reporting structure is through the clinical trial office administration. The more informal reporting structure is through a disease team function. These disease teams are often led by physician scientists and are responsible for the execution of clinical trials. Research staff are mediated by organization-wide controls (e.g. cancer center and/or university administration requirements) as well as team-determined goals. Oftentimes, dedicated staff members from the regulatory and clinical research coordination/nursing teams are assigned to work with specific disease teams. This design is very similar to the combined business unit (BU) and scientific knowledge base units that are described by Griffith et al. (41). In their framework, the combination of team members with differing expertise proved to be more productive as it supports knowledge sharing between team members with different types and depth of knowledge, improving performance (67). The disease teams are not only responsible for executing clinical research, but proposing new clinical research to activate. Investigators obtain clinical trial protocols, either from industry, the NCI National Clinical Trial Network (NCTN), through academic collaborations, or by writing the protocol themselves. It is the disease team that initially promotes activation of the clinical trial protocol, through formal or informal processes, depending on the institution.

2.5.1.1 Multiple Team Membership (MTM)

Several papers discuss the positivity in productivity when centralized multi-disciplinary teams and/or multitasking in the science field are utilized (41, 43). The Multiple Team Membership (MTM) framework is described as the relationship between staff productivity and the number of projects a team member is handling (42). This is presented as an inverse curvilinear relationship with lower productivity with lower team/project interaction and increasing to a point where the person becomes overwhelmed, at which point productivity decreases (*ibid*). This framework has been applied to the R&D field and demonstrated that there is an optimal point which balances multitasking by a team member and maximum output. Applying this to clinical research, where staff often juggle multiple clinical research studies, this framework can be utilized to examine what is the optimal workload for staff working on study activation, as well as the organizational context of the types of projects overseen by the team member, which was highlighted by Blindenbach-Driessen as an important context in cross-functional teams (44).

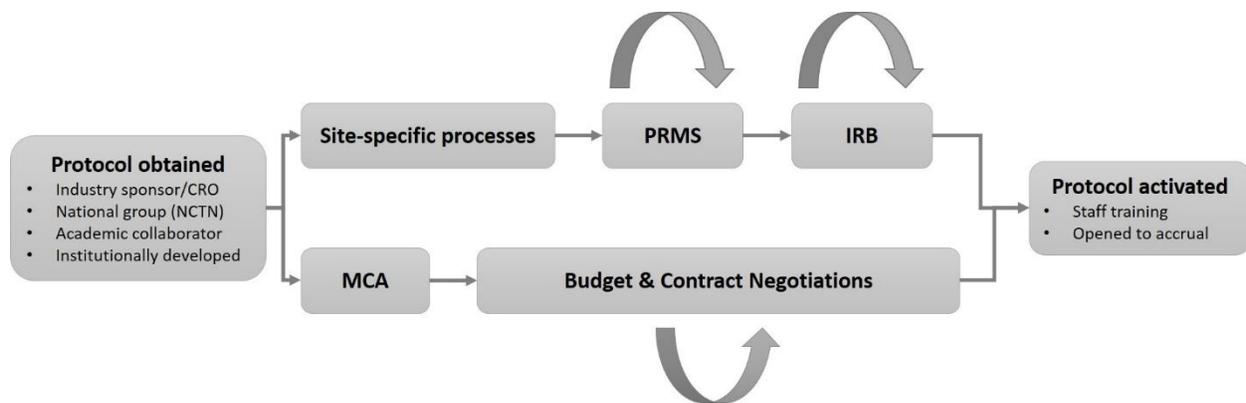
2.5.2 Scientific Review within the NCI-Designated Cancer Center System

One requirement for NCI-designated cancer centers which perform clinical research is to have a scientific review mechanism, designated as the Protocol Review and Monitoring System (PRMS). This is an additional and distinct committee from the IRB with a mandate to review the scientific merit of a clinical research protocol as well as monitor the accrual progress of ongoing protocols (68). Applying the stage-gate framework, these committees act as gates in the study activation process, forming a 3-stage/2-gate process to open a study: pre-PRMS stage, PRMS gate, pre-IRB stage, IRB gate, study activation. While this adds an additional step in the study

activation pipeline, the purpose of this review mechanism is to ensure that only protocols which are scientifically meritorious and relevant are opened at an NCI-designated cancer center. Clinical research undergoing formal external peer review (such as those sponsored by the National Clinical Trial Network [NCTN] or funded through a federal agency) receive an administrative review at the PRMS stage, rather than review by the full committee. Therefore, the majority of protocols reviewed by this mechanism are industry- or institutionally-sponsored. A simplified flow of the clinical trial activation process at an NCI-designated cancer center is visualized in **Figure 2-2**.

Figure 2-2: Simplified Study Activation Pipeline

High level overview of the processes and milestones at the site level to activate a clinical trial.



Ning et al. examined the impact of this extra committee requirement at a single NCI-designated cancer center. Their analysis found that 40% of trials received an outright approval by the committee. All other protocols required some level of revision and 1% were disapproved and not continued. Revisions were both protocol-related and non-protocol related, falling into the

categories of: evidence/rationale for the study, study design, intervention, population, consent form, and other. Institutional protocols were twice as likely to require revisions, and had an average of 5.6 changes per protocol versus 2.4 changes for industry protocols (69). Early phase trials (phase I/pilot or phase II) had an initial approval rate of approximately half that of phase III studies, a statistically significant finding. While the mandate for the PRMS is to be separate from and not overlap with the IRB, this group found that 51% of requested changes were related to the consent form and other non-protocol documentation, potentially constituting mission creep of the committee (*ibid*). This group did not evaluate the accrual outcomes of these protocols undergoing review to correlate whether this review improved ultimate accrual to the study. However, if only 1% of protocols were disapproved, it appears that protocol feasibility was not a primary focus of this committee, providing support for the argument that another group with this focus is needed.

2.5.2.1 Feasibility Committees

While both the PRMS and IRB committees review the protocol at an NCI-designated cancer center, neither is mandated to perform a comprehensive review of whether the staffing and infrastructure capacity exist to properly execute the trial in a timely fashion. To address feasibility, some sites have implemented an administrative review to assess staffing and/or workload capability, called in this project a “feasibility review.” The purpose of this committee is to avoid opening protocols that will not succeed. If resources are thoughtfully considered during a feasibility review, then studies that are unlikely to accrue any subjects should never be opened, or if they are, opened rarely. The decision to create and manage a feasibility review

committee has, to date, been left to the institution; however, institutions have instituted the feasibility review process early on in the study activation timeline to maximize the savings of staff resources for protocols that ultimately will not activate. While, in theory, the addition of another formal review process increases the length of time to activate a study, it should result in the center having a higher accrual of subjects per protocol. However, this has not been examined in the peer-reviewed literature. Applying the stage-gate framework, the addition of a formal feasibility process adds another gate to the process, creating a 5-stage/3-gate process: pre-feasibility stage, feasibility gate, pre-PRMS stage, PRMS gate, pre-IRB stage, IRB gate, and, finally, study activation.

2.6 Clinical Trial Accrual

It has been well reported that many studies do not meet their national accrual goals (16). This has been a documented issue since the 1970s, when accrual to clinical trial was discussed during the NIH-hosted 1978 National Conference on Clinical Trial Methodology (70). Since that conference papers every decade have highlighted, especially in the cancer setting, the high number of trials that do not accrue the pre-determined number of subjects to address the questions they set out to answer. Denny et al. reported in 2004 that 9.4% of CALGB protocols from 1989-1997 did not meet accrual goals (22). A 2008 survey of study chairs and/or lead statisticians reported that 41% of survey respondents “indicated that their trial experienced significant accrual difficulties” (71). The respondents represented 223 cancer cooperative group trials conducted between 1993 and 2002. Of these studies, 37% were classified as having insufficient accrual, with both therapeutic trials (those targeted at testing an intervention directed

at the person's cancer) and non-therapeutic trials (which can include supportive care and prevention trials), having the same proportion of trials classified as having insufficient accrual. Reporting on phase III cooperative group studies conducted between 1998 and 2002, Schroen et al. found that 52.5% of studies closed due to insufficient accrual (23). Several studies reported on the rate of under-accruing studies in the 2000s. Between 2000 and 2004, approximately 38% of Cancer Therapy Evaluation Program (CTEP) trials, the NCI program that vets NCI-sponsored clinical trials—including 70% of phase III trials—did not meet accrual goals (24). The IOM reported that 40% of phase III cooperative group protocols opened between 2000 and 2007 did not meet accrual goals (2). Dilts et al. reported that 45% of trials from a sample of all participating NCI cooperative groups during the same timeframe closed with insufficient accrual (18). In the same year, Dilts et al. reported that 6.4% of Cancer Therapy Evaluation Program (CTEP) trials did not accrue a single subject nationally (*ibid*). Recently, a report from the American Society of Clinical Oncology (ASCO) reported that 20% of all phase II and III cancer clinical trials in ClinicalTrials.gov registered between 2005 and 2011 did not finish for reasons unrelated to the treatment or procedure. Low accrual was the number one reason for early discontinuation of a trial (72).

Sites often do not meet accrual goals for multi-site clinical trials. The percentage of studies that do not accrue any studies at the site-level has been reported from approximately 20% up to 50% (1, 18, 25-26). Over three-quarters (77.4%) of cooperative group trials accrued less than five subjects at the site-level while 45.4% of non-cooperative group trials accrued less than five subjects (18). The paper examined processes associated with the development and activation of a

cooperative group trial and, under that context, the authors suggest that the lengthy time it takes to open such a trial impacts the ultimate accrual to trial, leading to an increased failure rate over industry-sponsored protocols. The number of layers of approval needed to open a cooperative group trial are more than industry trials, due to NCI-mandated processes. In addition, the authors discuss the lack of financial reimbursement for accruing subjects to cooperative group trials which, at the time of this publication, was \$2,000, limiting the ability and/or desire of institutions to enroll subjects onto such trials.

Not obtaining the needed number of research subjects during the dedicated recruitment period leads to prolonged recruitment phases. The IOM reports that clinical trial enrollment periods are estimated to be extended in approximately 90% of trials worldwide (1). For example, a review of NIH-sponsored HIV/AIDS studies between 2006 and 2011 showed that phase III/IV therapeutic trials had the longest development time and took the longest to enroll with a median development time of 2.5 years and median enrollment period of 3 years (73). This leads to an increase in the number of sites and staff needed to conduct clinical research, inflating costs (3). This is also an issue of public health, as it delays the availability of safe and effective treatments to the general public (28).

If recruitment is not prolonged, then the trial ends incomplete, potentially affecting the power of the study. Low accrual leads to underpowered studies, meaning that the scientific question at hand is not properly addressed, leading to research waste (27). This also affects the ethical conduct of the study as, if no conclusions regarding benefit can be reached, then the risk-benefit

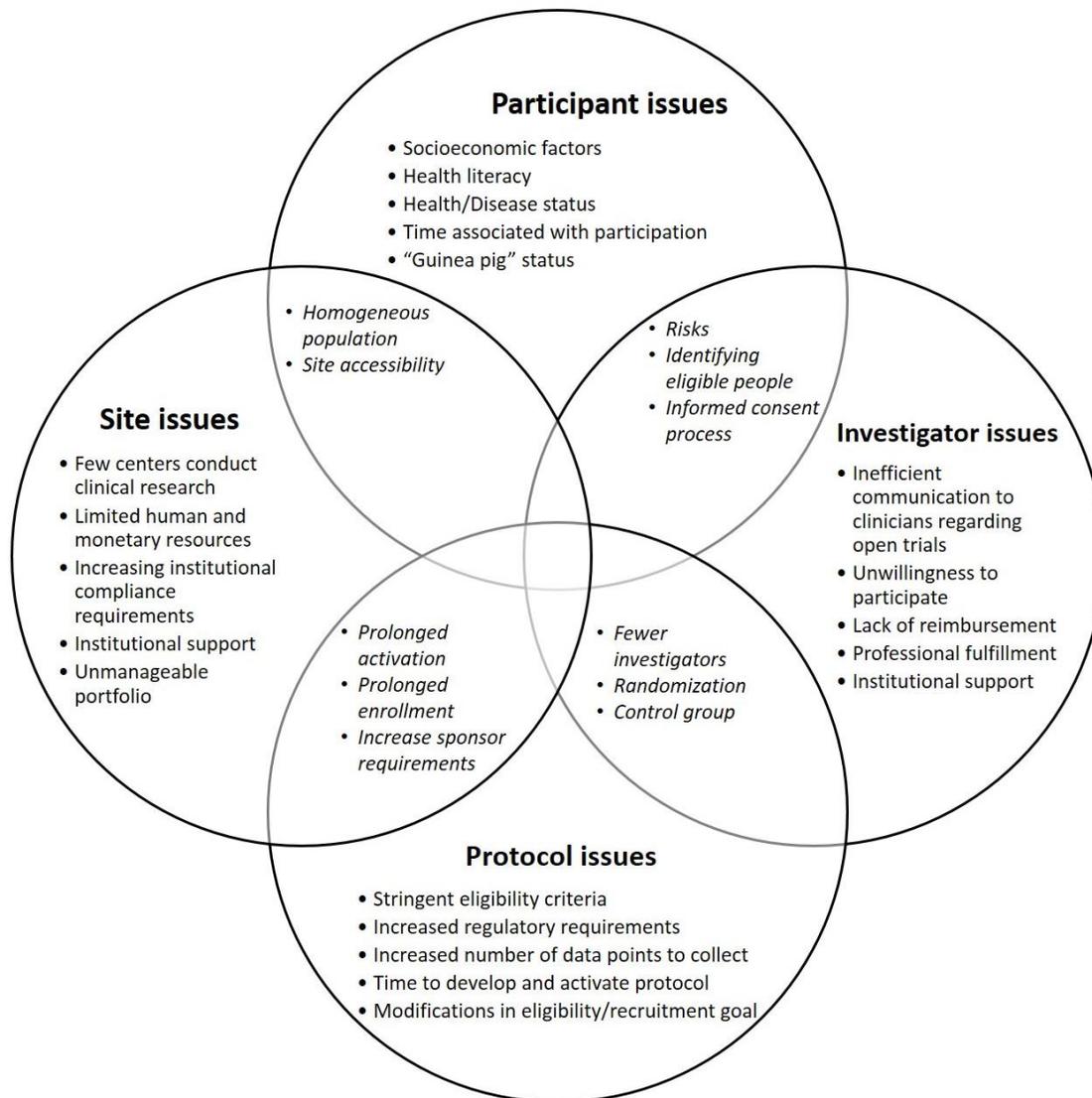
ratio is weighted towards risk. This threatens the Belmont principle of beneficence (74) and does not meet the required U.S. Department of Health and Human Services (DHHS) regulation that risks to subjects be minimized and risks are reasonable in relation to anticipated benefits [(45 CFR 46.111(a)(1) and (2)]. The result is that many valuable scientific questions regarding disease treatment go unanswered, as well as ongoing ethical debate of whether it was appropriate to place the subjects who did enroll in these trials at risk, as no benefit (the gain of scientific evidence for safety and/or effectiveness) was received.

2.6.1 Factors Affecting Clinical Trial Accrual

Low accrual is a widespread issue, affecting research in many disease states. Throughout the literature, issues of accrual in many diseases are explored. Cancer is commonly addressed, with additional literature regarding low accrual in cardiovascular disease, HIV/AIDS, emergency research, pediatric non-cancer disease, and neurocognitive conditions. A review of the literature showed that many of the publications regarding clinical trial accrual were reports from single clinical trials reporting the characteristics of the recruited population. Many of these studies were large clinical trials, such as the Program on Surgical Control of the Hyperlipidemias (POSCH; 75), Breast Cancer Prevention Trial (BCPT; 76), and the Cardiac Arrhythmia Suppression Trial (CAST; 77) and had separate sub-studies built into the protocol to assess factors associated with clinical trial enrollment. However, these studies sought to explain, post-hoc, pitfalls associated with enrollment specific to the trial and did not generalize accrual issues for application to other, future trials.

For this dissertation, a review of the literature was undertaken and studies discussing issues associated with clinical trial accrual were categorized into one of four overarching categories that affect clinical trial accrual. The four general categories, diagrammed in **Figure 1-1**, that impact clinical trial accrual are: protocol issues, site issues, investigator issues, and participant issues.

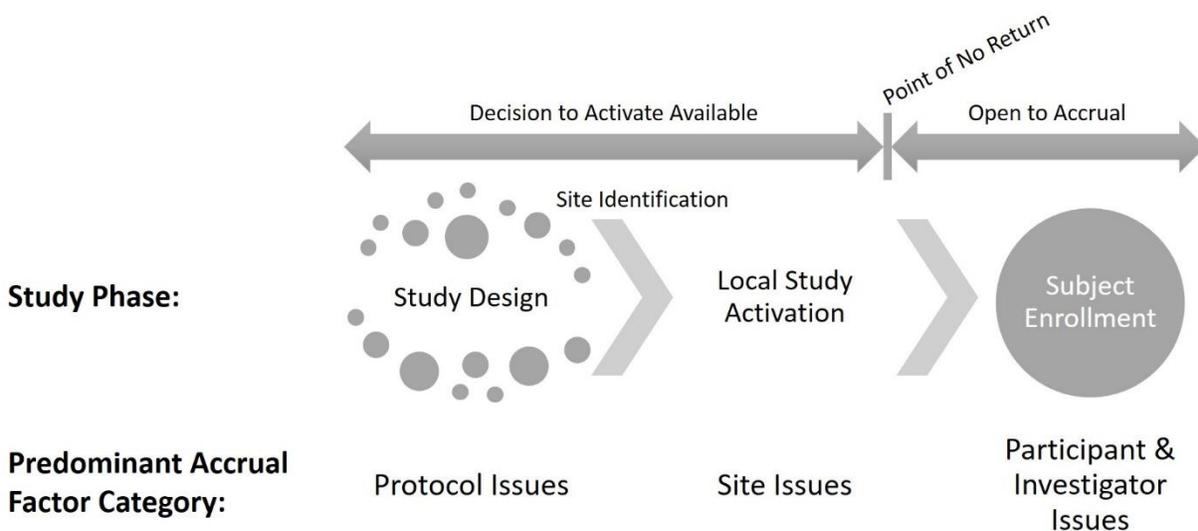
Figure 1-1: Factors Affecting Clinical Trial Accrual



These categories are not mutually exclusive. Issues affecting clinical trial accrual often fit multiple categories. For example, meeting eligibility criteria is both a protocol and participant issue. It is a protocol issue in that the investigators designing the protocol have the ability to modify the eligibility criteria to make it more or less restrictive, as scientifically appropriate. Meeting eligibility criteria is also a participant issue as potential subjects must meet all the criteria to be eligible. Stricter criteria limit the available pool of people from which one can recruit. It is important to note that these issues affect clinical trial accrual differently at various stages of study development, activation, and execution (**Figure 2-1**).

Figure 2-1: Factors Affecting Clinical Trial Accrual Considering Time

Factors affect clinical trial accrual at different stages of the clinical trial lifecycle. Diagramed below is the clinical trial accrual lifecycle with the accrual factor category that primarily affects clinical trial accrual at that stage.



Prior to study activation and during study design and development, protocol issues (such as inclusion/exclusion criteria and procedure selection) drive the ultimate ability to accrue subjects (78). After the protocol has been developed, the local site activation process begins. During this time site issues, such as the ability for a site to activate a study in a timely manner, the interest of the disease team (which overlaps with investigator issues), and having the needed staff and infrastructure in place to accrue subjects are most impactful on local accrual. After the local “point of no return,” opening the study to accrual, the investment in the study has been made. At this point investigator issues and participant characteristics are the driving factors for accrual success. Thus it is crucial to make choices that maximize accrual early on in the activation process to make the greatest impact. Focus should not be on factors that affect accrual after a study has been activated at a center, mainly participant and investigator issues, but on maximizing potential accrual prior to a study opening through understanding and choosing studies that are most likely to succeed.

2.6.1.1 Protocol Issues

Study design directly impacts how a protocol is executed, including site-level accrual (71). The complexity of clinical trials is increasing, with more inclusion/exclusion criteria, procedures, biomarkers, biopsies, etc. (10). In recent years, an increase in regulatory requirements (such as conflict of interest and privacy review) has been observed (3, 29). The increased complexity makes it more difficult to complete the clinical trial, highlighted by the large amount of data being collected, including more sensitive adverse event criteria, which leads to a need for more staff resources per trial and can detract participants from enrolling (30-31). If no increase in

research staff is allotted, then fewer trials are able to be opened. Additionally, stringent eligibility criteria simplifies the trial design and statistical analysis but increases recruitment difficulty, leading to a higher probability that the accrual goal will not be met (1, 21, 30).

Extensive examinations of the effect of protocol factors on clinical trial accrual have been published since the 1980s. These include phase (79-81), disease type (81), use of placebo (82), interventional type (80), interventional timing (82), randomization (80), study endpoints (2), study design (80, 83), study purpose (80), sample size (80), sponsor type (80), and the number of participating sites (83). In 1983, participant recruitment to the Coronary Primary Prevention Trial, a multicenter community clinical trial looking at coronary heart disease, reported its findings. A few years later, in 1987, the multisite POSCH trial attributed many issues associated with the rate of study accrual to protocol-related factors. Factors described by one or both studies these (now three) decades ago include:

- Elongation of the recruitment period (37, 75)
- Changes in the recruitment methods and procedures (37)
- Modification of protocol eligibility requirements (75)
- Change in the target number of subjects to be enrolled (75)
- Expansion of the number of sites needed to accrue the needed number of subjects (75).

Over the past 30 years these concerns have not changed. In an analysis performed by the CTSA Evaluation Key Function Committee in 2014, only 54% of studies reported meeting their local accrual goals and 50% of studies had longer recruitment periods than expected (84). A review

by Bernardez-Pereira et al. of cardiovascular studies from 2000 through January 2013 assessed factors associated with a study that was terminated due to low accrual (80). Approximately 11% of the 6,279 clinical trials were terminated early. Of the terminated trials, over half were terminated due to lower than expected recruitment. Factors leading to lower accrual rates and early termination included university/hospital funding of the study and mixed sources of funding. Behavior/diet interventions and single arm studies were statistically less likely to terminate due to low recruitment.

In late 2015, Bennette et al. published a logistic model to predict low accrual at the national level for NCI-sponsored National Clinical Trial Network (NCTN) trials (85). Their definition of low accrual was a trial accruing at less than 50% of target. Factors found to be significantly different between studies that accrued and those that did not accrue successfully included the number of competing trials, radiotherapy intervention modality, targeted therapy, new investigational agent, annual incidence of eligible patient population, enrollment fraction, randomized design, phase III studies, number of interventions studied, and more than one condition evaluated. This model is unique in that it utilized over 800 studies to determine which factors impact accrual, and then built and validated a model that had 97% fit in predicting whether a trial would accrue at least 50% of the stated goal (*ibid*).

Issues categorized as both protocol factors and participant factors were also identified. These include issues surrounding accessibility to the clinical research site (1), randomization

procedures (36), use of a placebo (36, 86), and an increased number of processes and procedures that are burdensome to participants (1).

Clinical trial development (design) time has been correlated with clinical trial accrual (24, 79). Cheng et al. showed that studies that took less than 12 months of development time had twice the odds of achieving their accrual goal than studies that took 12-18 months to develop. Studies that took greater than 24 months to develop were significantly less likely to meet their accrual goal than studies taking 12-18 months of development time (24). It was postulated that this could be due to many reasons, including the speed of which oncology treatments progress, thus making the ongoing clinical trial (at best) uninteresting or (at worst) obsolete; or due to a lack of scientific interest to dedicate significant resources to conducting the research (*ibid*). Once opened, the time to enroll the first subject at a site has been correlated to ultimate accrual (79, 87-88), with longer times associated with lower overall accrual, as has the accrual rate within the first few months of activation (88). These studies conclude that delays in getting pertinent scientific studies activated lead to underperforming accrual. This leads to waste within the clinical trial enterprise in the form of both monetary and human resources.

However, protocol design elements were not always perceived as determining factors in accrual. A survey of cooperative group study chairs and statisticians found that no elements of trial design (use of placebo, time between trial concept and activation, or eligibility criteria) were a factor in clinical trial accrual efficiency (71). A review of over 6,000 cardiovascular clinical trials found that most protocol factors (phase, endpoint, sample size, age, etc.) were not

associated with early termination due to low accrual (80). These conflicting results may indicate that there are other factors, perhaps in one of the other areas (participant, site, or investigator) or in the characteristics of the trial (e.g. disease area) that influence accrual at a greater factor, or that those factors modify the effect of these factors.

An assessment of cooperative group study chairs and biostatisticians viewed a cooperative group's experience in a specific disease, disease stage, or intervention as the strongest factors in predicting overall clinical trial accrual (71). Another study associated clinical trial accrual with sponsor communication methods and choice of data capture method (83). Buonansegna et al. developed a framework of why clinical trials fail from the sponsor perspective. Their industry analysis found seven management issues that led to the failure of clinical trials. These seven issues, targeted at the sponsor-side of clinical trial management included one protocol specific item: lack of feasibility of the study protocol. Sponsors who do not test the procedures to ensure they are practical and economic at the site level increase the protocol's chance of failure (31).

2.6.1.2 Site Issues

Several studies have found that both center type and center size are associated with accrual (83, 89-92). Most U.S. community practices and health systems do not incorporate research at their sites (29). It is reported that only approximately 5% of the approximate 5,000 U.S. acute care hospitals consistently participate in clinical trials (93) and it is estimated that two-thirds of sites never do more than a single trial (32). This means that much of the clinical research performed is done at academic medical centers. This is seen in cancer with only approximately 30% of

accrual to NCI-sponsored cancer treatment trials between 1998 and 2008 from community institutions; the other seven-tenths of accrual came from academic medical centers (94). Another report lists approximately half of clinical trial accruals as coming from academic centers (33).

An analysis of clinical trials show that a large proportion of trials are sponsored by academic health centers and are focused on disease mechanisms (34). Conducting such studies requires dedicated human resources. Additionally, as clinical trials become more complicated (21, 29), specialized staff and/or training is needed to execute clinical research efficiently and in a compliant manner. However, many affiliate hospitals and clinics of academic medical centers provide clinical researchers little support, and oftentimes the relationship between the research unit and the clinic is adversarial (93). Thus leading academic centers are disinclined to provide institutional support to fund the activities of investigators involved in industry-sponsored projects (*ibid*). In the same article, Califf et al. provide five suggestions to make the clinical research enterprise a more functional component of the learning health care system. These suggestions encompass modification to site funding, administrative streamlining, cultural changes regarding the view of clinical research, and network building. Every suggestion requires change at the national level, specifically the NIH (*ibid*).

Califf et al. highlighted the multiple layers of university bureaucracy, stating “academic health and science systems are hard-pressed to find efficiencies, and the clinical enterprise has little tolerance for slowing practice to accommodate prolonged consent processes or for deferring profitable procedures” (93). This is highlighted in the prevalence of the relative value unit

(RVU) system for measuring clinician work effort within academic systems and the prioritization of clinical practice over clinical research investment. In addition to the lack of institutional financial support of clinical trials, industry sponsors and clinical research organizations (CROs) increase financial difficulties as payments to sites for procedures and staff time already spent are, on average, delayed 120 days, leading to the average U.S. site carrying a clinical research debt of \$400,000 (95).

In the REACT study, institutional factors that affected clinical trial enrollment included the presence and activity of research staff. In particular, having one physician responsible for enrollment (versus rotating enrollment responsibilities) and nurse coordinators involved in the study were associated with increased site enrollment to the CAST trial. Other studies explored the involvement of the research staff in the recruitment process. Often higher engagement and/or enthusiasm regarding study participation by health care/research personnel led to enrollment to clinical trial (77, 92, 96-98). Additionally, staff (particularly research nurses) positively influenced a person's decision to join a clinical trial (92). A study exploring factors of effective clinical research teams in cystic fibrosis showed, as determined through a survey ranking system, that factors that impacted clinical trial accrual include multi-level shared leadership (e.g. principal investigator and research coordinator both having leadership roles), a customer service orientation towards participants, institutional culture promoting research, communication between the clinical team and research team, and staff longevity (99).

The REACT study also found that an affiliation with a medical school and having house staff and cardiology fellowship training programs were associated with higher enrollment. This partners with sponsor selection and finding highly qualified investigators. The sponsor-choosing sites who do not have the proper facilities, infrastructure, and personnel can lead to the failure of a clinical trial, thus reducing the pool of potential subjects to recruit from (31).

Like the other sections, site factors are influenced by the other types of factors. The presence (or absence) of protocols competing for similar patient populations affect accrual (81, 89). Protocol factors that affect site-level accrual include effective study design, prioritization, and accrual prediction practices (71). The FDA-sponsored interdisciplinary expert panel also raised concerns about clinical trials, specifically that patient populations are not always representative; that “proven therapies” are not always provided as part of the clinical trial; costs of providing therapeutic intervention, follow-up and evaluation is not always appropriate; and disease characteristics (progress and epidemiology) are not representative (100). However, even if a study is designed well and the above site factors are addressed, it is up to the site and the site investigator to choose studies that will accrue at that site.

Sponsors impact the conduct of sites in clinical research. Current ability and characteristic surveys distributed by sponsors to sites that assess the ability to enroll subjects to trial do not correlate with subsequent accrual (100). The conduct of the site initiation visit of a sponsor impacts the ability to accrue to a clinical trial (83). In Buonansegna’s framework of why clinical trials fail from a sponsor perspective, several findings were associated with site factors. Chaotic

and slow patient recruitment was a main finding. Delays in identifying the ideal patient population immediately after a product is chosen for testing, as well as sites approaching potential participants in a biased fashion, lead to study failure. This could be perceived as an investigator issue and/or site issue. In addition, lack of experience in choosing and monitoring partners was a factor. The sponsor-choosing sites who do not have the proper facilities, infrastructure, and personnel can lead to the failure of a clinical trial (31). Likewise, sites need to ensure that they can meet the requirements outlined by sponsors prior to agreeing to take on a clinical trial. Another issue is with data quality. Mistakes in clinical trial documents drastically influences the outcomes associated with the study product.

The most relevant factor to this dissertation is the unmanageable level of portfolio complexity. Too many protocols makes it difficult to manage all protocols, the sites, and associated requirements, leading to clinical trial failure. Finally, at this sponsor level, but also applicable to the site level, is the impact of incorrect assessment of the market potential or returns. When expenditures exceed the return on investment, the product can be discontinued, categorizing the clinical trials associated with the product as “failed” (31).

2.6.1.3 Investigator Issues

Overall, there is a decline in number of U.S. investigators with an increase in clinical trials worldwide and global outsourcing of clinical trial work (1, 21, 34). The IOM reports that, as of 2007, 85% of investigators only ever conduct one clinical trial (1). The reduction in experienced investigators means there is more competition for sponsors to find highly-qualified investigators

to conduct clinical research, especially in rare disease areas (10). Another report extends this difficulty to other critical research staff, such as biostatisticians and clinical informaticians (68).

A survey of all ECOG investigators was performed in 1988 to assess factors associated between investigators and their accrual to clinical trial (101). After the survey, the accrual of each participating investigator was tracked. During this one-year period, 62% of investigators did not enroll any subjects to an ECOG clinical trial. Ten percent of investigators accrued 80% of subjects; 10% also accrued at least five subjects to trial. Investigator characteristics that led to higher accrual included a more research-oriented view of clinical research (versus viewing research as a treatment option), principal investigator status, and medical oncologist specialty (101). Contrary to the REACT sub-study of CAST, this study showed that physician rank was not associated with accrual (98, 101). Several groups have studied physicians associated with the community cancer organization CCOP, along with organizational characteristics that affect cancer clinical trial enrollment (89, 91, 102). The CCOP was made up of nonacademic oncology practices which were heavily involved in NCI-sponsored research. In 2011, when the study took place, approximately 40% of CCOP physicians enrolled no patients; the mean was five patients during the study period. Factors that contributed to higher enrollment include being a CCOP principal investigator, years of tenure/practice, more positive attitudes towards clinical trials, and supportive policies and procedures surrounding research. Negative factors included age and a non-oncology specialty (e.g. gynecology, surgery). The study found that physician factors contributed more to accrual than site factors (102).

Investigator issues and site issues are closely related, as the site employs the investigator. An expert panel of clinicians, researchers, sponsors, and regulators within the cardiovascular field met in 2014 as part of an FDA-sponsored meeting and noted several investigator-related challenges (also separately documented in peer-reviewed literature) with conducting clinical trials (100). Reports regarding infrastructure issues surrounding investigator participation in clinical research include work goal requirements (e.g. relative value unit [RVU] requirements; 30, 100). It takes a significant amount of work and time to monitor ongoing subjects on clinical trial (29, 103). Unless an investigator gains authorship on the trial publication, this is often without any performance credit for the intensity of that work to the investigator; only enrollment of a subject gives credit to the investigator (30, 100). A lack of funding at the institutional, sponsor, and federal levels discourage clinicians from more actively engaging in clinical research (10, 21, 27, 100). Federal funds limit salary support for investigators (100). Sponsors increasingly delegate clinical trial oversight and regulatory responsibility to clinical research organizations (CROs), which decreases available funding to sites and investigators (10, 100). A lack of discretionary funds to the investigator to conduct research of interest and to offset salary support also discourages investigators from being more involved in research (100).

Protocol factors also influence investigator impact on clinical trial accrual. A survey of cooperative group study chairs and statisticians reported that sufficient clinical trial accrual was associated with the perceived clinical relevance of the clinical trial by the investigator, the pragmatism of the trial and its applicability to clinical practice, and a lack of competing trials (71). Another study concluded that investigators who do not optimize recruitment tactics to

recruiting the appropriate people to trial in a timely manner contribute to the ultimate failure of a clinical trial (31). In his paper regarding theoretical frameworks regarding clinical trial participation decision-making, Morrow suggested that physician factors are the greatest reason for non-enrollment and that, by targeting physicians, clinical trial accrual can be increased (104).

Finally, investigators are often asked to predict how many people they will accrue to trial. This prediction is used in industry-sponsored feasibility questionnaires, clinical trial contracting, and often PRMS review. However, research suggests that investigators are over-optimistic in their accrual estimations. Our own analysis concurred with this finding, with investigators overestimating accrual by 70-300% (105). ECOG investigators overestimated accrual by a factor of six. Overestimation was not more pronounced in non-oncologists than oncologists (101). Thus, sites can be set up for failure from the beginning solely due to the investigator's optimism on predicted accrual.

2.6.2 Modeling Clinical Trial Accrual

Multiple regression modeling is suited for predicting accrual as accrual is a multi-dimensional outcome with several contributing parameters that can be used to calculate an outcome. In this case, the outcome of interest is the predicted number of accruals for a given clinical trial. The representative format for multiple regression is:

$$y = \beta_0 + \beta_1 * x_1 + \dots + \beta_n * x_n + \varepsilon$$

where y is the value of the dependent variable (outcome), β_0 is the intercept, β_i is the magnitude of change associated with the corresponding independent variable (x_i), x_i is the value of the independent variable, and ε is the error.

For this study many of the outcomes of interests—clinical trial accrual, particularly—are not a normally distributed, continuous variable, but positive count data which is highly right-skewed. Regression analysis utilizing a negative binomial distribution (a special case of the Poisson distribution) accounts and adjusts for this non-normal distribution. The calculated values from each independent variable and its associated weight (beta value) are summed and then exponentiated on a log₂ scale, giving a value for the dependent variable (outcome). The applicability of this method to accrual prediction was highlighted in a statement by London et al. who wrote “the trick is... getting [researchers] to recognize the uncertainty inherent in the accrual rates and quantify it in terms of a prior distribution” (106). The prior distribution is represented by the regression equation. Regression modeling also allows for the use of a logistic distribution, which will be used for binary outcomes.

During a review of the literature, 29 items were identified that focused on either the statistical methodology or computational modeling of clinical trial accrual. Of the 29 items, 19 were peer-reviewed publications (107-125), three were patents describing computational displays for clinical trial calculations (126-128), two were systematic reviews of the literature (47, 129), two were dissertations (one focused on screening techniques and the other focused on calculating accrual rates; 130, 131, respectively), one was a book chapter (132), one was a letter to the editor

in response to a peer-reviewed publication (133), and one was a conference abstract (134).

Twenty papers, the letter to the editor, the conference abstract, and the book chapter discuss statistical modeling and were based on utilizing regression modeling to calculate the accrual rate and/or forecast when a protocol would achieve its accrual goal (107-125, 131-134). It is notable that both Bayesian and Frequentist techniques were used to predict accrual. Eight papers using Frequentist statistics focused on accrual using a Poisson distribution (111, 115, 117, 120, 122-123, 131, 133) and two utilized Brownian motion (110, 118). It is important to note that these papers utilize the beginning of the recruitment period to predict the remaining time and accrual number; it is not truly predicting accrual prior to study start. Additionally, trial-level accrual does not occur at a linear rate; accrual occurs stochastically, with a slow start, exponential phase, and then tapers towards the end (113). This is measured by these papers, but may not be reflective of what happens at the site level. Five papers utilized Bayesian techniques (107, 119, 121, 124-125); however, these also focused on trials where accrual was ongoing.

An accrual model fits the framework presented by Barnard for the use of predictive models in that it would be:

- 1) Simple to use and understand (the outcome is a single number with actual meaning in the clinical trial enterprise)
- 2) Able to adapt to epidemiologic changes (through the modification of independent variables)
- 3) Able to adapt to environmental changes (through the modification of independent variables)

- 4) Able to take account of center recruitment
- 5) Able to inform commissioning decisions (47).

2.6.2.1 Previous Accrual Prediction Work

Having identified at our center that low accrual was a human and financial resource burden, we sought to create and validate a local predictive model with the outcome of anticipated accrual to be used when considering a prospective clinical trial. To do this, we conducted a retrospective cohort study using 5.8 years of registry data from treatment and supportive care interventional studies at our center for model construction. A negative binomial regression model was employed using variables known pre-study abstracted from the OnCore® clinical trial management system (Forte Research Systems, Inc., Madison, WI) and ClinicalTrials.gov. Accrual was predicted for studies used to build the model.

The model included 207 trials with complete information. Mean accrual was 7.3 per trial (± 18.4); 55 (26.6%) trials accrued zero subjects locally. In univariate analysis use of an investigational drug, disease team, number of national sites, use of local IRB, number of total months open nationally, months of accrual already completed, and overall proposed national enrollment were significantly associated with accrual. In multivariate analysis disease team, proposed national enrollment, number of sites, use of local IRB, number of total months open, and number of months already opened were significantly and independently associated with accrual. The full model was significant ($p < 0.001$) and predicted accrual at 94% of actual, maintaining predictive value at multiple cutoff values (105).

The validation study included 2.5 months of registry data using criteria from the original model. Studies were run through the prediction model and actual accrual plotted against predicted accrual. Actual team- and model-predicted subjects accrued, percent of trials meeting cut-off values, and model sensitivity and specificity were calculated. For the validation study, 61 trials met the inclusion criteria. Total accrual was 373 subjects (mean: 6.1 ± 17.2); 16 (26.2%) had zero accrual, 23 (37.7%) accrued 88.7% of the total subjects. The model predicted accrual of 513 subjects (138% of actual) versus the disease team predicted accrual of 1,111 subjects (298% of actual). The model correctly predicted whether a study would accrue at least four subjects 75.4% of the time. Twenty-seven studies (44.3%) correlated perfectly at the category level. Model sensitivity is 70.0%; specificity is 78.1%. For the 17 studies not correctly categorized using a cutoff of four, nine (60%) would have been wrongly opened (predicted 4+, <4 accrued) and six (40%) would have incorrectly not opened (predicted <4, 4+ accrued).

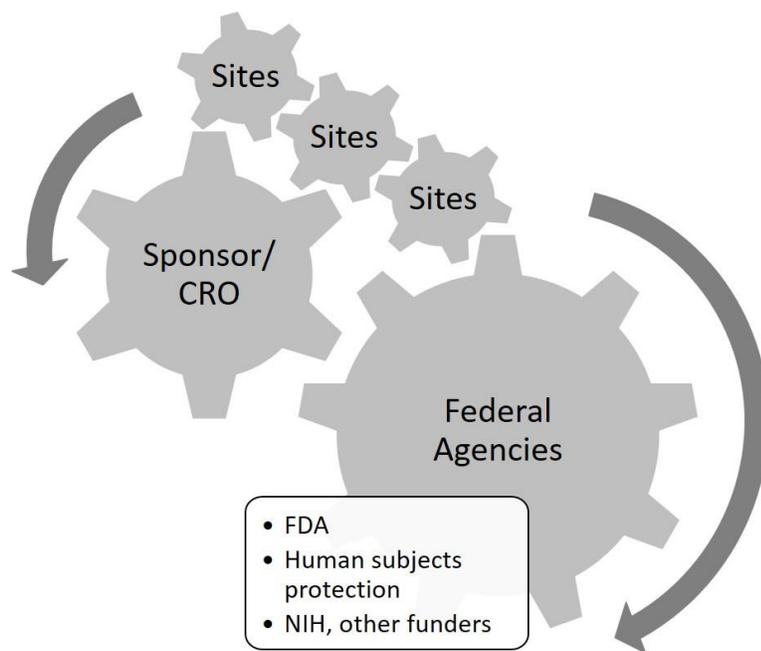
We identified key factors, both nationally and locally, associated with subject trial accrual at our site. This model can aid in deciding whether a study is likely to accrue a desirable number of subjects. The validation of this model shows it to be an accurate, quick, and valuable metric in assessing trial success as well as planning resource allocation and clinical trial costs. This work was presented as a poster at both the 2014 and 2015 American Society of Clinical Oncology (ASCO) annual meetings and a peer-review publication is forthcoming.

2.7 Gaps in the Literature

The clinical research enterprise is like a machine made up of cogs (**Figure 2-3**).

Figure 2-3: Key Clinical Trial Players

Stakeholders within the clinical research enterprise can be described as cogs that interlink into a complete working machine. The size of the cog indicates its national impact on clinical research operations.



Each cog impacts the rate at which the machine turns; however, the amount of impact depends on how much the one cog can influence the overall clinical research process. Players that have great impact across many factors of the clinical research process, such as regulatory agencies which set national policy and requirements to conduct research, are larger cogs in the clinical research machine. Smaller players, such as sites, also impact the clinical research enterprise (as they are the reason that clinical trials accrue subjects); however, a change at the single site level is unlikely to greatly affect the entire clinical research machine. Coordinating several smaller

cogs, however, can make a large impact, turning the larger wheels of the machine. Current research is focused on modifying the larger cogs, the clinical trial activation process at the national level (with sponsors and/or federal regulators overseeing clinical research). While this approach will have the broadest impact, it takes significant time and energy to make such changes. Despite that fact this is where the energy of modifying the clinical research enterprise has taken place. However, if focus can be redirected towards coordinating the smaller cogs and making changes at the site level, the influence on the overall clinical trial enterprise will be positively impacted without national policy change.

Despite national consortia examining issues surrounding clinical research management, there remains a gap in knowledge regarding what constitutes the proper workload, or staff to protocol ratio, to most efficiently activate a clinical trial. As there are no national standards for clinical trial office staffing, institutions must develop their own process organically and several different staffing models exist, particularly around regulatory management and study activation. Thus it is **unknown** whether staff devoted to study activation are more efficient at opening studies that will ultimately accrue participants than team structures that have more diverse job descriptions encompassing multiple aspects of clinical research conduct.

Regarding processes and methodologies, there is a gap in knowledge within the clinical trial space of when **not** to open, rather than opening, a clinical trial at a site, investing human, physical, and financial resources, not meeting accrual goals, and ultimately closing the trial at the site. Not all trials are made for equal performance at every site. There is a lack of

methodologies and/or tools available to determine which trials are not going to succeed before study activation.

Several types of recommendations exist in the peer-reviewed literature to improve accrual; however, these did not include any business processes or mathematical modeling for objectively determining whether a study should be activated (30, 100, 104). Given the number of studies stated above that indicate that many trials accrue poorly at the site level, and the lack of literature regarding quantitative and objective measures to calculate clinical trial accrual *a priori*, it appears that there is a gap in knowledge on **how** to get researchers to recognize the inherent uncertainty that Gajewski describes (107) and objectively calculate a realistic prediction of subjects they are likely to accrue. This search of the literature shows that several models have been created to assist in determining who will specifically join a clinical trial, whether a trial will meet its accrual goal based on its predicted accrual rate, and whether the trial as a whole will accrue the needed number of subjects. It also discusses factors that lead to low accrual, including participant factors, investigator issues, protocol issues and site issues. However, while many studies look at individual site accrual, it may be so site-specific that the external validity to other clinical trial sites is limited. On the other hand, factors found to affect accrual at the national level may not accurately reflect accrual barriers experienced at individual sites (71). These are topics not explored in the current peer-reviewed literature.

The focus of the issues associated with study activation and accrual were based on studies that were already activated and open to accrual (10). This is a more reactive, rather than proactive,

approach and studies that will not accrue will still be activated and opened for enrollment. It would be optimal (for both the sponsor and site) to not invest the extensive human and financial resources to activate a study at the site level that will not accrue a desired number of subjects. As it takes a large amount of time and money to initiate a research site—some industry reports cite the cost as \$50,000 per site—not activating low accruing sites can save money at a rapid rate without much additional work (personal correspondence).

2.8 Summary

This work addresses concerns raised by the IOM and independent researchers regarding research waste associated with the length of time associated with the clinical research process. By streamlining activation processes at the site level and activating studies at a site that will be successful in accruing subjects, the clinical development process is expedited, reducing research waste and improving efficiency. To address the current gaps in knowledge regarding clinical trial activation and accrual, four specific aims are proposed that apply industrial-organization (I-O) management frameworks and statistical modeling to clinical research, giving sites a process and tools to objectively assess whether a clinical trial should be activated. Developing this approach will address documented IOM, sponsor, and researcher concerns regarding the high percentage of sites that do not accrue subjects to trial. Implementation of this process at sites will maximize clinical trial efficiency by dedicating human and financial resources towards studies that will contribute to the mission of the center, impacting science and its patient population. This work impacts health policy by providing a method to maximize the timely

accrual of subjects to clinical trials, progressing scientific knowledge and providing high quality information to inform clinical practice.

CHAPTER 3. METHODOLOGY

3.1 Study Purpose

The purpose of this dissertation project is to address the gap in knowledge regarding how to objectively describe and quantitate optimal study activation workload at the study site level while maximizing the probability of clinical trial success, defined as site clinical trial accrual.

3.2 Study Design

This is a retrospective cohort study of cancer centers utilizing administrative data previously collected as part of clinical trial operations. The level of measure (unit) in this study is the protocol. Outcomes (e.g. time to activate a study, total accrual) are specific to this unit. Thus, the inclusion of multiple institutions creates a hierarchical structure with factors that are specific to a protocol and independent of other protocols regardless of institution, as well as institutional factors that will influence protocols within that institution but not between institutions. Mixed regression modeling will be utilized with a random effect (the institution) to account for this hierarchical structure.

3.2.1 Disease Teams

Within a cancer center, infrastructure is further designated by disease teams. Disease team is the generic term for an organized group of investigators conducting research. Within cancer centers, these teams are often classified by disease site, but can also be cross-disciplined and organized by another factor, such as pediatric oncology or early phase trials (two common multidisciplinary classifications). The defining members of these teams are the investigators, clinician-researchers

who are involved in the scientific development of institutionally sponsored (and possibly externally sponsored) protocols, determination of interest in taking responsibility for a clinical trial and shepherding the protocol through the activation process, maintaining the protocol once opened (including the identification, consent, and oversight of subjects), and ensuring the data is collected and submitted to the sponsor as outlined by the protocol. Under federal regulations (both the Office of Human Research Protections and the Food and Drug Administration [FDA]), the investigator who takes on overall oversight of a protocol at the site (the principal investigator) is held responsible for all conduct of the trial at the site. Due to the great responsibility associated with being a principal investigator, clinician-researchers and disease teams weigh this decision carefully. Clinical trial personnel may be hired by or assigned to (in the case of a centralized office) a specific disease team, adding additional infrastructure and clinical trial support. With these trial responsibilities, disease teams directly affect the ability to accrue to a clinical trial as well as are an influencing factor in how quickly a trial can be activated.

3.3 Center Recruitment

3.3.1 Inclusion Criteria

Centers are eligible for inclusion if they were NCI-designated (either clinical or comprehensive) as of January 1, 2015. There is an exception to this criterion, as two institutions that did not have NCI designation but followed the NCI designation guidelines expressed interest in participation and were allowed to submit data for this project. Studies were eligible for inclusion if they were 1) conducted at a participating center; 2) therapeutic or supportive care, interventional protocols

(as defined by the NCI Data Table Guide, version July 29, 2013; 135); and 3) whose local status permanently moved to “closed to accrual” after January 1, 2009. Expanded access, treatment IND, and single-patient use protocols were not eligible for study inclusion as these are not research projects (per the FDA) but treatment options. Per the Cancer Center Support Grant (CCSG) Data Table Guide, interventional studies are defined as studies where “individuals are assigned prospectively by an investigator based on a protocol to receive specific interventions. The participants may receive diagnostic, treatment, behavioral, or other types of interventions. The assignment of the intervention may or may not be random. The participants are followed and biomedical and/or health outcomes are assessed” (135). Supportive care protocols are “designed to evaluate one or more interventions where the primary intent is to maximize comfort, minimize side effects, or mitigate against a decline in the participant’s health or function. In general, supportive care interventions are not intended to cure a disease” and treatment protocols are “designed to evaluate one or more interventions for treating a disease, syndrome, or condition. Note: This equates to therapeutic trials in previous versions of the guidelines” (*ibid*).

3.3.2 Recruitment Methods

NCI-designated centers were identified from the NCI website for cancer centers (136). Emails explaining the project and offering participation in this project were sent to cancer center administrators (**Appendix A**). After three non-responses (spaced approximately one month apart from the previous correspondence), centers were considered non-responsive and no further contact was made. Centers responding positively regarding participation were sent a

standardized introduction letter and data abstraction sheet in Excel 2010 (**Appendices B and C**).

Centers submitting data before December 15, 2015, were included as participating centers.

3.4 Data Collection

3.4.1 Cancer Center Protocol Data Collection

Center-specific study characteristics for all eligible studies were abstracted from the local clinical trial management system (CTMS) by local cancer center clinical trial administrators or their designee. Center-specific characteristics abstracted from the clinical trial management system included protocol identifier, ClinicalTrials.gov identifier (NCT ID, primary key), disease team (categorical), center-specific (non-affiliate) actual accrual (continuous), disease team predicted accrual (continuous), use of local IRB (yes/no), cooperative group sponsor (yes/no), the date the study started the activation process locally (numerical), the date the study was opened to accrual locally (numerical), and the date the study closed to accrual locally (numerical). Disease team was cleaned to keep only the primary team who oversaw the study.

Protocol-specific characteristics were abstracted from ClinicalTrials.gov, a U.S. federal government website which lists all clinical trials contributing to an approval application with the Food and Drug Administration (FDA), trying to publish in an International Committee of Medical Journal Editors (ICMJE) journal, or by voluntary submission outside these requirements. These protocol-wide characteristics included sponsor name, phase (categorical), randomized design (yes/no), use of placebo/clinical observation (yes/no), primary and secondary outcome measures/endpoints (safety/toxicity only, efficacy, or other), inclusion of pediatric

subjects (yes/no), national enrollment goal (continuous), the number of participating sites (numerical or unknown, as applicable), the national study start date (numerical), and the national primary endpoint completion date (numerical). The data collection sheet is found in **Appendix B** and the data dictionary for the full dataset in **Appendix C**.

3.4.2 Cancer Center Administrative Data Collection

Participating centers were asked to complete a data collection form describing center infrastructure for the study activation process (**Appendix D**). Information regarding staff resources dedicated to regulatory processes, both full-time equivalent (FTE) and number of staff, was requested. Classification factors, such as whether staff were housed within the centralized clinical trials office, whether staff were dedicated to only performing regulatory functions or if they also acted as clinical research coordinators, and if they only worked on study activation were obtained. To capture changes over time in staffing levels, the status of each of these factors was requested at six month intervals starting in January 2009 and ending in January 2015.

In addition to staffing information, data regarding activation was obtained. Center-specific workflows for study activation by sponsor type were ascertained via data collection form and then follow-up email and/or phone interviews, as necessary for clarification. Protocols with “New” date (date when the study activation process began) were associated with the administrative infrastructure that was present at the institution at that time.

3.4.3 Data Integrity

3.4.3.1 Data Cleaning

Submitted protocols that had one or more of the following factors missing were removed from the analysis: NCT ID, Local Open Date, Local Close Date, National Open Date, National Close Date, or Local Accrual. Submitted data was checked for face validity. Protocols were removed if the ClinicalTrials.gov listing categorized the protocols as anything other than “Interventional” (i.e. observational or expanded access), the ClinicalTrials.gov listing categorized the primary purpose as something other than “Treatment” or “Supportive Care” (e.g. Health Services Research, Basic Science), if the protocol was not an oncology protocol, the local number of days open to accrual was less than or equal to zero, if the national recruitment time was less than zero (the listed Local Close Date was before the Local Open Date), if the Local Open Date was after the National Close Date, or if the Local New Date was after the Local Open Date. For studies where the Local Open Date was before the National Open Date, it was assumed that the Local Open Date was the National Open Date as the institutional local to accrual date is the more reliable source of information. Prior to removing the protocols from the data set, the institution was contacted for clarification. Where possible, data was updated based on clarifications/corrections made by the institution.

3.4.3.2 Variable Definitions and Manipulations

A number of variables were not available within the clinical trial management systems or ClinicalTrials.gov. Thus manipulations of the collected data generated the final dataset that was used to analyze the specific aims, sub-aims, and hypotheses of this project.

For each six-month period, an indicator of whether an internal feasibility committee for studies was used (yes/no), and regulatory staffing (both number of staff members and FTEs) generated from administrative data collected from the site. For institutions providing data regarding the use of a feasibility committee, it was assumed that protocols opened after the date corresponding with the implementation of the feasibility committee utilized the committee (e.g. if a feasibility committee was implemented in January 2014, then all protocols with a local open to enrollment date of January 2014 or later were assumed to utilize the committee).

For the purposes of this project, the inclusion of placebo was defined as a protocol having at least one of the following scenarios: 1) an observational arm (defined as no intervention AND not receiving usual/standard of care); 2) any use of a placebo, including arms that included an active drug in addition to placebo; or 3) mention of placebo as a treatment. This definition of placebo was utilized as the inclusion of any of these arms could influence the participation of a potential subject, thus affecting accrual.

During abstraction of the number of national sites from ClinicalTrials.gov, it was found that centers were not consistently included in study records, and in other records a center would be listed multiple times (e.g. separate campuses of the same institution). Therefore this variable was categorized into five categories to best describe the size of the study. These were: 1-9 sites, 10-49 sites, 50-199 sites, 200+ sites, and number of sites unknown. The inclusion of the category “unknown sites” is relevant for predictive purposes as many studies (particularly

national group protocols) do not indicate the intended number of sites that are planned to contribute to the enrollment goal.

From the ClinicalTrials.gov record, three additional categorical variables were created. The first was whether a protocol was a precision medicine trial. Precision medicine, for the purposes of this project, was defined as a protocol having at least one inclusion or exclusion criterion in the ClinicalTrials.gov record that included at least one specific genetic aberration (mutation, etc.) determined by sequencing (not histology, immunohistochemistry, cytology, karyotyping, etc.). Studies were classified as being not precision medicine, precision medicine, or conditional precision medicine. Conditional precision medicine studies were those studies where at least one of the following criterion were present: 1) if mutational status is known, it must be disclosed but does not disqualify a person from participating; 2) the mutational status does exclude if known, but testing is not required to participate; or 3) the mutational status does include/exclude but is one of a list of criteria that are annotated by an “or” and not all the criteria are mutation-based.

The second variable created was sponsor type. The Study Source categories from the Cancer Center Support Grant (CCSG) Data Table guidelines were used as the basis to define sponsor type (135). Sponsor type was determined from the sponsor listed on the ClinicalTrials.gov record. Categories include:

- National Group: NCI National Clinical Trials Network (NCTN) or other NIH-supported national trial network
 - Sponsor was an NIH entity and the primary investigator was from an NIH entity

- Sponsor was a non-profit consortium with an NIH collaborator
- Externally Peer Reviewed: Sponsor was an academic institution and a national funder (e.g. NIH or other federally funded clinical research mechanism) listed as a collaborator
 - Studies who listed the NIH or other federally funded clinical research mechanism as the sponsor, but the primary investigator was from an academic institution
- Institutional: Sponsor was an academic institution and no federal funder was listed as a collaborator
- Industry: Sponsor was a for-profit company, regardless of listed collaborators

Institutions provided data regarding whether the study was classified as a National Group study or not (binary variable). These values were compared to the ClinicalTrials.gov abstraction and where discrepancies existed, more information regarding the sponsor, participating collaborators, and location of the principal investigators was obtained to ensure that studies were categorized according to the above definitions.

Finally, the national closure date was abstracted from protocols that were no longer open to enrollment. If the actual date was provided, then this date was used. Where only month and year were provided, the day of the month was set to the first. If the actual date was not provided, then the latest release date of the record changing the status of the study from a recruiting status to a non-recruiting state was utilized. Where the projected/actual National Close Date was not available, the national primary endpoint completion date was used as a proxy for the date that accrual closed nationally. This field, defined by ClinicalTrials.gov, is “the date that the final subject was examined or received an intervention for the purposes of final collection of data for

the primary outcome, whether the clinical trial concluded according to the pre-specified protocol or was terminated (137).” This proxy value is appropriate as the length of time to accrue the total number of subjects to protocol and the time to obtain the primary outcome are relative to each other. They differ by a constant variable, the length of the intervention which, mathematically in a regression model, is a negligible value. Additionally, it is estimated that the enrollment period in 90% of clinical trials is extended (1), which means that it is not unreasonable to assume that the recruitment period is closer to the anticipated primary completion date than the planned recruitment end date.

From the collected data, additional variables were created. These included whether the protocol was a zero-accruing protocol (yes/no), whether the study accrued at least four subjects (yes/no), the months of accrual completed, calculated as the number of days between the National Open Date and the Local Open Date divided by 30.4, and the total enrollment period, calculated as the number of days between the national start date and the national date of expected primary endpoint completion divided by 30.4. The magnitude of difference between the disease team accrual prediction and actual accrual was generated. For studies where a disease team prediction was present, but local accrual was zero, a value of 250 was set, as this was larger than the largest actual difference observed of 210.

The distribution of each continuous variable was assessed for normality and transformed as necessary. The transformation matrix for each continuous variable is shown in **Figure 3-1**.

Figure 3-1: Continuous Variable Transformation Matrix

Figure 3-1a: Months of Accrual Completed

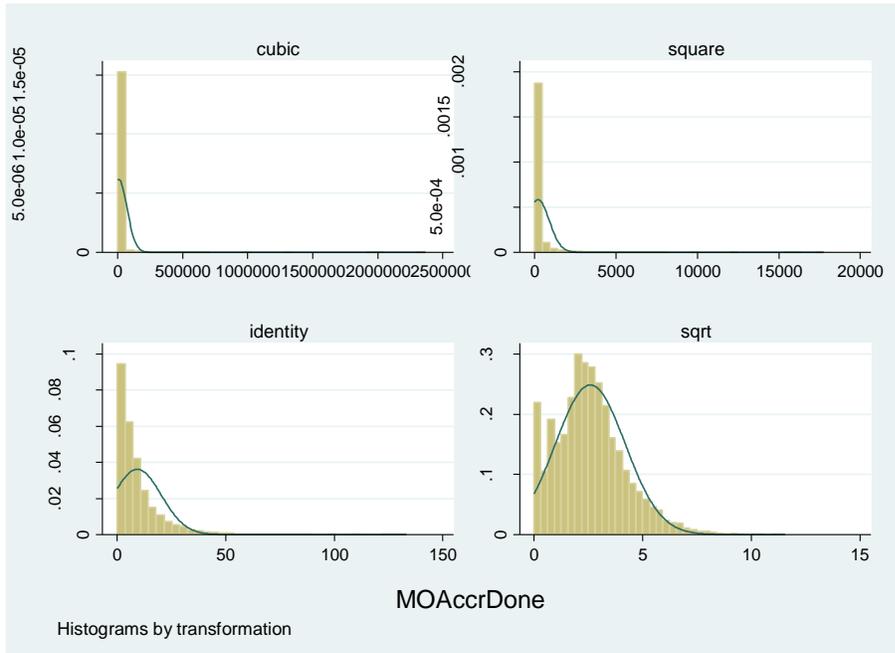


Figure 3-1b: National Enrollment Goal

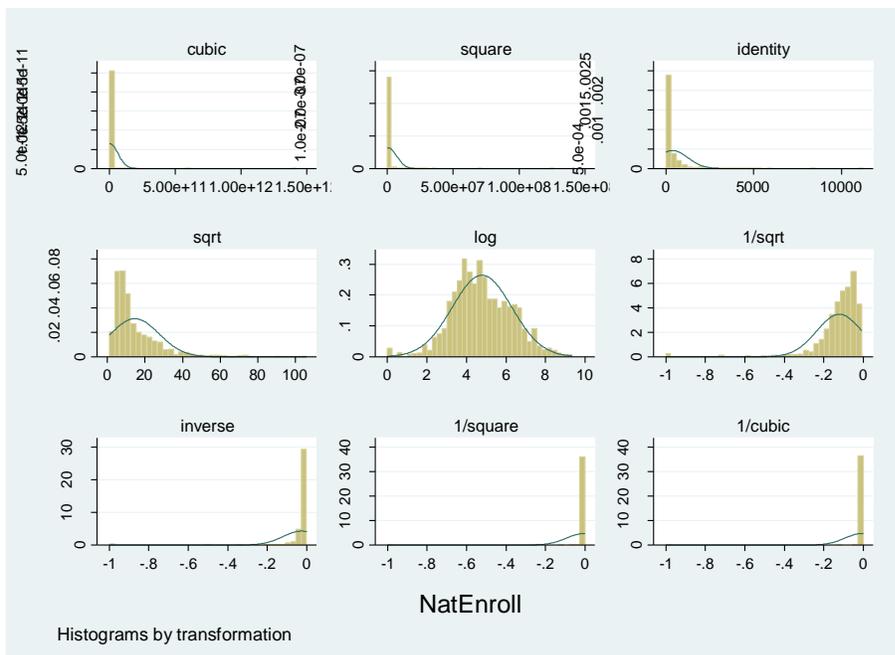


Figure 3-1c: Total Months of Accrual

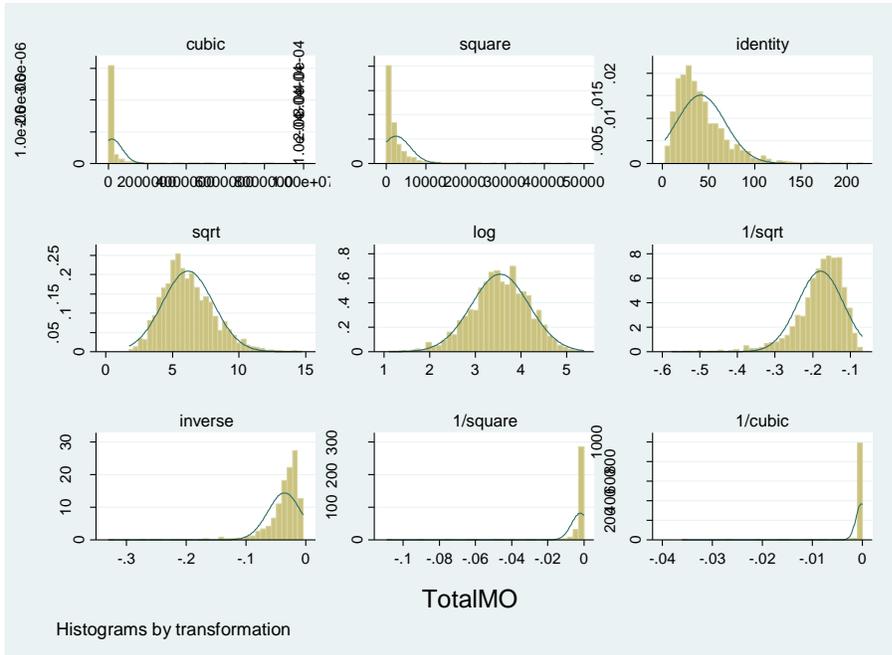
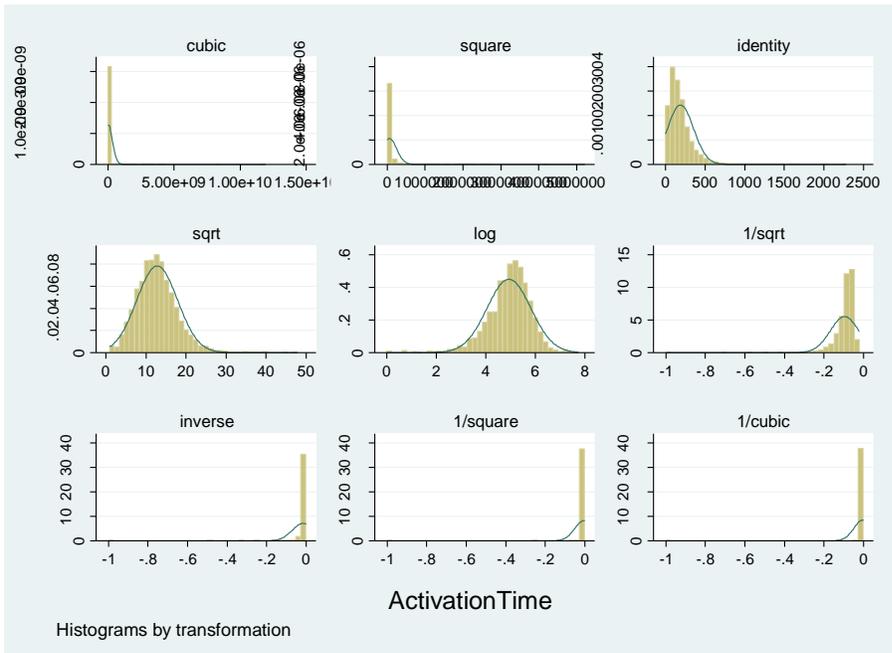


Figure 3-1d: Activation Time



3.4.3.3 Calculation of Protocol per Staffing Unit

For each institution, the number of protocols starting the activation process (indicated by the “New” status date provided by the institution) was calculated for six month intervals (activation period). Then the number of protocols per activation period was divided by the staffing unit provided by the institution (number of staff members or number of full-time equivalent [FTE] positions). For institutions providing only FTE levels and indicating that the staff is dedicated to regulatory work, it was assumed that the number of staff members equaled the number of FTEs.

3.5 Data Analysis

3.5.1 General Statistical Considerations

All statistical tests utilized a significance level of $p < 0.05$. Summary statistics, including frequencies, averages, standard deviations, medians, and ranges were generated for each variable to describe the participating centers as well as the protocols utilized in the analyses.

Stratifications of summary statistics were calculated for the following categories: accrual (all protocols, zero-accruing protocols, and protocols accruing at least one subject), 3-stage/2-gate process (no feasibility committee) or 5-stage/3-gate process (feasibility committee), and protocols contributing staffing information for Specific Aim 2.

Differences in categorical variables were assessed using chi-square tests (tests containing cells with under five observations utilized the Fisher’s exact chi-square test). Differences in continuous variables were tested using t-tests assuming unequal variances (in comparing

continuous variables in two categories) or ANOVA (for continuous variables with more than two categories).

3.5.2 Regression Analyses

3.5.2.1 Distributions

The outcome of accrual is a positive count that is highly right-skewed, thus a negative binomial distribution was utilized in regression analyses.

The outcome of zero-accruing protocol is a binary value (yes/no), thus a logistic distribution was utilized in regression analyses.

The outcome of study activation time is a continuous value, thus a linear distribution was utilized in regression analyses.

The listing of dependent/independent variable pairings is listed in **Table 3-1**.

Table 3-1: Dependent-Independent Variable Pairings

DEPENDENT VARIABLE	INDEPENDENT VARIABLE	SPECIFIC AIM
1. ACTUAL ACCRUAL	Feasibility committee (5-stage/3-gate versus 3-stage/2-gate)	Hypothesis 1a
2. PERCENT ZERO-ACCRUING STUDIES		Hypothesis 1b
1. STUDY ACTIVATION TIME	Number of Protocols per Staff Member	Hypothesis 2a
2. ACTUAL ACCRUAL		Hypothesis 2b
3. PERCENT ZERO-ACCRUING STUDIES		Hypothesis 2c
1. STUDY ACTIVATION TIME	Number of Protocols per FTE	Hypothesis 2d
2. ACTUAL ACCRUAL		Hypothesis 2e
3. PERCENT ZERO-ACCRUING STUDIES		Hypothesis 2f
1. PREDICTED ACCRUAL	All independent variables to be tested for association	Sub-Aims 3a & 3b

3.5.2.2 General Regression Analysis Considerations

In regression analyses that involve multiple institutions a random effects model, using institution as the random effect, was used to account for the hierarchical nature of the data. Institution was picked as the random effect as each center has infrastructure that independently (from other centers) and consistently (within the center) affects the conduct of unit of measure, the protocol.

As to not overfit the model, the number of variables included in the full model was no more than 15% of the limiting sample size. Limiting sample size was determined as outlined by Harrell (138). For continuous and count outcomes (clinical trial accrual and study activation time), the limiting sample size is equal to the total sample size. For binary outcomes (zero-accruing protocols), the limiting sample size is equal to smaller sub-population size.

Measures of fit, such as pseudo- R^2 , likelihood ration (LR), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) values were utilized to assess which model best fits the data. In a non-linear model, the pseudo- R^2 statistic cannot be interpreted as the R^2 statistic is in the linear model, as it cannot describe the amount of the outcome explained by the model. For the purposes of this project, pseudo- R^2 values were used when constructing models as an aid to assess what model was more complete (with higher values being desired). When constructing models, the AIC and BIC was considered to distinguish which model is a better fit of the data. For both tests, a lower value indicates better model fit. A likelihood-ratio test testing the model fit of the negative binomial versus a Poisson was run on regression models. A statistically

significant value (<0.05) indicates that the model run was a better fit than the same model using the Poisson distribution.

3.5.2.3 Reduced Models

Univariate analyses were run using the distribution listed above (in **Section 3.5.2.1 Distributions**) specific to the outcome of interest to measure the association of each independent variable with the dependent variable (**Table 3-1**; page 115). Independent variables included various site, protocol, and investigator factors that were collected from each participating site's clinical trials management system as well as ClinicalTrials.gov. Reduced models testing for association were performed to test for effect modification using a threshold of modification set at 10% change in the beta coefficient. Interaction models for variables known to affect national enrollment (i.e. number of national sites, total months planned for national accrual, number of months of national enrollment already completed, and sponsor type) were run on reduced models found to have an effect modification. Interactions found to be statistically significant at a p-value of less than 0.05 were assessed as part of the full model.

3.5.2.4 Full Models

Factors found to be significant ($p<0.05$) in the univariate models and the interaction models were integrated into the full model to control for their association with the dependent variable. An unadjusted model using the appropriate distribution for the outcome of interest and associated covariables were run to assess, irrespective of institution, whether the outcome differs based on

the independent variable of interest. Adjusted models were also run, utilizing a random effects mixed model, with institution as the random effect.

Full models were generated using all significant variables and interaction combinations. The best model with the best AIC and BIC scores were examined for fit based on degrees of freedom. Every attempt was made to keep the number of degrees of freedom used to 10% of the total number of observations; however, in site models with smaller numbers of observations ($n < 200$), the models produced with this constraint were not well fit. Therefore, the strongest variables were included until the number of degrees of freedom used was no more than 15% of the number of total observations. Fifteen percent of the total number of observations was chosen as it helps assure a reliable, fitted model, according to Harrell et al. (138).

3.5.3 Protocol Workload Sensitivity Analyses

To assess the hypothesized inverse curvilinear relationship between outcome and protocol workload (Specific Aim 2), several combinations of workload and staffing models were considered. Staffing models include number of full-time equivalence (FTE) and number of staff members. Outcomes to measure efficiency were 1) overall accrual, 2) percent zero-accruing protocols, and 3) time to activation.

As one inclusion criterion for study inclusion was a requirement that the protocol be closed to accrual, a site's entire protocol activation workload portfolio was not available. To calculate total activation workload, only protocols opened from 2007 through 2010 were assessed. These

years were chosen as the median number of months that a protocol is accruing nationally is 35 (average 42) with three-quarters of protocols closing to accrual nationally within 54 months. The average time a study is open to accrual at a site is 26 months (median 21 months), with 75% of protocols closing in under 36 months and 90% of protocols closing to local accrual within 3.5 years. Thus, by using the years 2007 through 2010, the majority of protocols that were opened would have closed by the beginning of 2015 when studies were submitted for this project, giving an approximation of total workload. To compensate for the protocols that accrue longer than the 75th-90th percentile, a sensitivity analysis was performed by increasing optimal protocol workloads by 10% and 25% and presenting the range.

3.5.4 Accrual Assessments

Accrual measures were categorized (<1, 1<4, 4<7, 7<10, 10<20, 20<50, 50+) and tabulated to assess what percentage of trials directly correlate between the two accrual measures being compared. Sensitivity, specificity, and overall accuracy were calculated to assess which model (the site-specific, unadjusted overall, or adjusted overall model) more precisely predicted actual accrual. Accuracy was calculated as the percentage of trials that were correctly predicted at a cut-off of four accruals. Site-specific model-predicted accrual and overall model-predicted accrual were compared to actual accrual using Pearson's correlation coefficient. Disease team-predicted accrual and the predicted accrual were compared to actual accrual using Pearson's correlation coefficient (Sub-Aims 3c and 3d).

3.6 Assumptions

A major assumption with this modeling approach includes that the institutional circumstances leading to accrual are consistent with future studies conducted by the institution, reminiscent of the infamous quote “past performance is the best predictor of future performance.” As national regulations and requirements regarding the conduct of clinical research are stagnant and many centers are slow to change institutional processes, this is a rational assumption for this analysis.

Another, more process-based assumption includes the time that a study was open to accrual locally. If a study was opened to accrual and then subsequently closed temporarily (e.g. for an interim analysis), the time the study was closed did not reduce the overall time a study was considered open to accrual for the purposes of this project.

3.7 Anticipated Limitations

Due to the retrospective nature of this study, there are potential missing factors that could improve the accuracy of this model but are not captured in the site’s clinical trial management system or ClinicalTrials.gov. Efforts have been made to extrapolate relevant factors from these databases, such as the delineation of precision medicine protocols and the actual national accrual closure date; however, it is not guaranteed that all contributing factors to clinical trial accrual are accounted for in the studied variables. There are also data integrity limitations. As data is being abstracted from the local clinical trial management system and the ClinicalTrials.gov database, the data is only as accurate as it has been entered.

Finally, this study is only assessing cancer centers and cancer protocols. Due to the federal requirements surrounding cancer center operations, as well as the complex nature of oncology clinical trials, this work may not be generalizable beyond a cancer center adhering to the NCI Cancer Center Support Grant (CCSG) deadlines.

3.8 Computational Support

Data abstraction was in Excel 2013 in comma-delimited format. Statistical analysis was performed using STATAIC 14 (STATA Corp., College Station, TX).

3.9 Confidentiality and Human Subjects

Participating cancer centers were coded with a six-digit code generated from a random number generator. This code will be used for identification purposes in all public presentations, reports, and publications regarding the characteristic summaries, regression models, and other statistical output. Each site will have knowledge of their code, but other participating sites will not have access to the code list. A list of participating centers are included in the dissertation paper and will be listed as participating sites in resulting peer-reviewed publications and conference presentations, unless otherwise specifically requested by the site to have the identity redacted. Members of the University of Arizona research team, specifically the principal investigator/PhD candidate (Wendy Tate) and her dissertation committee, will have access to the identifiable information for the purposes of data analysis and quality assurance.

The information collected as part of this project is either publically available and/or not about a living individual. Therefore, it does not meet the definition of a human subject, as defined in 45 CFR 46.102(f) and IRB approval is not required.

CHAPTER 4. RESULTS

4.1 Center Recruitment

NCI-designated centers (n=61) were identified from the NCI website for cancer centers (136). All clinical cancer center (n=20) and comprehensive cancer center (n=41) websites were reviewed for clinical trial office administration contact information. Contact information for all but one comprehensive cancer center was obtained. Recruitment emails containing a description of the project and presentation materials from the American Association of Cancer Institutes Clinical Research Initiative (AACI-CRI) annual meeting (**Appendix A**) were sent to the email addresses of clinical trial office administrators found. Institutions who did not respond were sent a reminder email approximately one and two months after the initial invitation. Institutions who responded negatively to the recruitment email were thanked for their consideration and not contacted again. Institutions who responded positively were sent the protocol data abstraction sheet, including instructions and a data dictionary (**Appendix B**). If an institution did not send in the abstraction sheet, a reminder of the sheet and inquiry on whether they had any questions was sent approximately every two to four weeks. The majority of centers provided initial protocol data prior to April 2015 with the last center providing data in December 2015 and updated data and clarifications provided no later than January 2016.

Fourteen centers (23%) submitted protocol information and were classified as participating centers. Two non-NCI designated centers also submitted protocol information. These centers asked to participate based on presentations heard at national conferences. Both centers adhere to the NCI Cancer Center Support Grant (CCSG) guidelines, requiring a Protocol Review and

Monitoring System (PRMS) process and utilizing a centralized clinical trials office, so they were allowed to participate. Therefore, a total of 16 centers participated in this project. **Figure 4-1** contains a map showing the geographic locations of the participating centers. These centers are (in alphabetical order by state):

- The University of Arizona Cancer Center; The University of Arizona, Tucson, Arizona (NCI comprehensive cancer center)
- Samuel Oschin Comprehensive Cancer Institute; Cedars-Sinai Medical Center, Los Angeles, California (Not NCI-designated)
- UCSF Helen Diller Comprehensive Cancer Center; University of California at San Francisco, San Francisco, California (NCI comprehensive cancer center)
- Stanford Cancer Institute; Stanford University, Stanford, California (NCI clinical cancer center)
- Moffitt Cancer Center, Tampa, Florida (NCI comprehensive cancer center)
- Robert H. Lurie Comprehensive Cancer Center; Northwestern University, Chicago, Illinois (NCI comprehensive cancer center)
- Indiana University Melvin and Bren Simon Cancer Center, Indianapolis, Indiana (NCI clinical cancer center)
- Holden Comprehensive Cancer Center; University of Iowa, Iowa City, Iowa (NCI comprehensive cancer center)
- University of Michigan Comprehensive Cancer Center, Ann Arbor, Michigan (NCI comprehensive cancer center)

- Masonic Cancer Center; University of Minnesota, Minneapolis, Minnesota (NCI comprehensive cancer center)
- Mayo Clinic Cancer Center, Rochester, Minnesota (NCI comprehensive cancer center)
- Alvin J. Siteman Cancer Center; Washington University School of Medicine and Barnes-Jewish Hospital, St. Louis, Missouri (NCI comprehensive cancer center)
- University of Cincinnati, Cincinnati, Ohio (Not NCI-designated)
- Fred and Pamela Buffett Cancer Center; University of Nebraska, Omaha, Nebraska (NCI clinical cancer center)
- Fred Hutchinson/University of Washington Cancer Consortium, Seattle, Washington (NCI comprehensive cancer center)
- University of Wisconsin Carbone Cancer Center, Madison, Wisconsin (NCI comprehensive cancer center)

Figure 4-1: Participating Centers

Centers contributing data to this project are marked with a black star



4.1.1 Infrastructure Data Collection

In Fall 2015, participating centers were asked to provide additional information regarding their infrastructure (to fulfill the analyses planned for Specific Aims 1 and 2). A separate invitation email and abstraction sheet (**Appendix C**) was sent to the center contact who provided the protocol information. Seven centers provided complete additional infrastructure information (regulatory staffing, protocol dates encompassing the activation timeline, and use of a feasibility committee). Centers providing complete infrastructure information are:

- Samuel Oschin Comprehensive Cancer Institute
- Stanford Cancer Institute
- Moffitt Cancer Center
- Robert H. Lurie Comprehensive Cancer Center
- Indiana University Melvin and Bren Simon Cancer Center
- Alvin J. Siteman Cancer Center
- University of Cincinnati

Not all institutions provided all the requested infrastructure information. The following institutions provided information regarding use of a feasibility committee, to define studies as following a 5-stage/3-gate process or a 3-stage/2-gate process:

- The University of Arizona Cancer Center
- Holden Comprehensive Cancer Center
- Fred Hutchinson/University of Washington Cancer Consortium
- University of Wisconsin Carbone Cancer Center

The following institution provided protocol dates encompassing the activation timeline:

- University of Wisconsin Carbone Cancer Center

The following institution provided regulatory staffing information:

- Fred Hutchinson/University of Washington Cancer Consortium

4.2 Summary Statistics

4.2.1 Institutional Summary Statistics

Institutional summary statistics are presented in **Table 4-1**. Eleven institutions (68.75%) were designated as NCI comprehensive cancer centers. Three (18.75%) were NCI clinical cancer centers and two (12.5%) were not NCI-designated. A total of 5,787 protocols were contributed by these 16 institutions for this study. Institutions contributed between 1.6% and 12.0% of studies, with the median percent contribution of studies being 5.95%, or 344 studies.

These 5,787 protocols accrued 49,319 subjects. Institutional contribution ranged from 0.7% of the total number of subjects (363 subjects) to 13.6% (6,705 subjects). Average accrual per trial was 8.5 subjects (standard deviation 17.5 subjects) with a median of four subjects accrued per trial. When evaluating accrual by site (versus aggregate), the median average accrual per protocol by institution was 9.6 subjects, with a range of 2.0 ± 3.5 subjects to 14.7 ± 26.7 subjects. Median accrual ranged from one subject per protocol to seven subjects per protocol. Removing the zero-accruing protocols, the average accrual increased to 10.4 subjects per protocol, median of five subjects (**Table 4-2**).

Table 4-1: Institution Summary Statistics

	# PROTOCOLS CONTRIBUTED	TOTAL ACCRUAL	AVERAGE ACCRUAL ±SD	MEDIAN ACCURAL	NCI DESIGNATION	TIMEFRAME SUBMITTED
ALL SITES	5787 (100.0%)	49319 (2.2%)	8.5 ± 17.5	4		
SITE 104647	365 (6.3%)	3497 (7.1%)	9.6 ± 16.0	5	Comprehensive	05/09-03/15
SITE 173472	407 (7.0%)	3407 (6.9%)	8.4 ± 15.5	3	Comprehensive	01/09-04/15
SITE 448155	697 (12.0%)	6716 (13.6%)	9.6 ± 14.2	4	Comprehensive	01/09-12/14
SITE 494048	552 (9.5%)	4029 (8.2%)	7.3 ± 14.9	3	Clinical	01/09-11/15
SITE 512786	484 (8.4%)	6228 (12.6%)	12.9 ± 26.5	7	Comprehensive	02/06-03/15*
SITE 560623	179 (3.1%)	363 (0.7%)	2.0 ± 3.5	1	Not Designated	07/08-11/15*
SITE 575415	410 (7.1%)	3549 (7.2%)	8.9 ± 14.4	4	Comprehensive	01/09-04/14
SITE 598430	326 (5.6%)	1338 (2.7%)	4.1 ± 6.9	2	Comprehensive	01/09-01/15
SITE 602591	395 (6.8%)	2113 (4.3%)	5.3 ± 7.5	3	Comprehensive	01/09-01/15
SITE 689326	456 (7.9%)	6705 (13.6%)	14.7 ± 26.7	7	Comprehensive	08/07-01/15*
SITE 696337	182 (3.1%)	831 (1.7%)	4.6 ± 5.6	2.5	Not Designated	06/09-11/15
SITE 714415	529 (9.1%)	5530 (11.2%)	10.5 ± 26.7	4	Clinical	11/08-02/15*
SITE 715532	225 (3.4%)	1313 (2.7%)	5.8 ± 10.8	2	Clinical	12/99-02/15*

SITE 846594	294 (5.1%)	1854 (3.8%)	6.3 ± 9.7	2	Comprehensive	01/09-01/15
SITE 997056	93 (1.6%)	754 (1.5%)	8.1 ± 13.4	4	Comprehensive	01/14-03/15
SITE 998666	193 (3.3%)	1092 (2.2%)	5.7 ± 8.9	2	Comprehensive	01/08-10/13
NATIONAL ENROLLMENT		2267084	391 ± 1567.7	104		

Timeframe submitted in the format of month and year. National enrollment based off of the data provided for each trial on ClinicalTrials.gov.

** Protocols with closure dates prior to January 2009 are incomplete (e.g. not all protocols for listed timeframe before January 2009 submitted).*

Table 4-2: Study Accrual Summary Statistics

	AVERAGE ± SD (MEDIAN)				
	TOTAL ACCRUAL	ACCRUAL – ALL PROTOCOLS	ACCRUAL – ACCURING PROTOCOLS	# MONTHS OPEN NATIONALLY PRIOR TO LOCAL ACCRUAL	# DAYS OPEN TO ACCRUAL LOCALLY
ALL SITES	49319	8.5 ± 17.5 (4)	10.4 ± 18.8 (5)	9.19 ± 11.03 (5.89)	791.6 ± 629.7 (627)
SITE 104647	3497	9.6 ± 16.0 (5)	11.0 ± 16.7 (5)	8.29 ± 11.34 (5.10)	638.9 ± 404.8 (543)
SITE 173472	3407	8.4 ± 15.5 (3)	10.1 ± 16.5 (5)	10.22 ± 11.62 (7.17)	798.5 ± 593.4 (664)
SITE 448155	6716	9.6 ± 14.2 (4)	11.2 ± 14.7 (6)	8.18 ± 9.77 (5.26)	782.6 ± 582.3 (634)

SITE 494048	4029	7.3 ± 14.9 (3)	9.3 ± 16.2 (4.5)	7.22 ± 10.33 (4.21)	810.0 ± 645.1 (640)
SITE 512786	6228	12.9 ± 26.5 (7)	14.2 ± 27.5 (8)	8.41 ± 11.21 (4.74)	608.9 ± 444.4 (516.5)
SITE 560623	363	2.0 ± 3.5 (1)	3.7 ± 4.0 (2)	12.90 ± 11.97 (8.62)	708.4 ± 480.6 (555)
SITE 575415	3549	8.9 ± 14.4 (4)	10.9 ± 15.4 (6)	9.19 ± 12.20 (5.92)	864.3 ± 755.1 (679.5)
SITE 598430	1338	4.1 ± 6.9 (2)	5.9 ± 7.6 (3)	11.08 ± 13.26 (7.38)	872.7 ± 694.4 (685.5)
SITE 602591	2113	5.3 ± 7.5 (3)	7.0 ± 7.9 (4)	9.96 ± 9.97 (7.11)	780.7 ± 590.6 (620)
SITE 689326	6705	14.7 ± 26.7 (7)	14.7 ± 26.8 (7)	7.37 ± 7.86 (5.28)	831.7 ± 599.2 (675.5)
SITE 696337	831	4.6 ± 5.6 (2.5)	5.8 ± 5.8 (4)	13.91 ± 11.69 (10.54)	683.1 ± 678.5 (463.5)
SITE 714415	5530	10.5 ± 26.7 (4)	12.3 ± 28.6	8.66 ± 11.12 (5.00)	835.1 ± 690.8 (639)
SITE 715532	1313	5.8 ± 10.8 (2)	8.3 ± 12.1 (4)	12.73 ± 13.46 (9.14)	977.7 ± 731.8 (806)
SITE 846594	1854	6.3 ± 9.7 (2)	8.1 ± 10.3 (3.5)	8.09 ± 10.37 (4.74)	998.1 ± 807.3 (773.5)
SITE 997056	754	8.1 ± 13.4 (4)	9.3 ± 14.0 (4)	10.56 ± 10.74 (7.50)	740.7 ± 597.1 (597)
SITE 998666	1092	5.7 ± 8.9 (2)	7.9 ± 9.7 (4)	9.77 ± 9.22 (7.63)	692.4 ± 573.3 (563)

4.2.1.1 Representativeness of Study Population

Sixty-one (61) institutions were NCI-designated as of January 1, 2015, when recruitment for this project initiated. Fourteen (14) institutions (23% of all institutions) ultimately provided information that met the inclusion criteria to participate in this study. An additional two centers (Samuel Oschin Comprehensive Institute and the University of Cincinnati Cancer Center) were not NCI-designated, but because they followed the NCI Cancer Center Support Grant (CCSG) guidelines in their cancer centers, their data was included for analysis. Of the 47 non-participating institutions, two expressed interest in participating but could not provide the requested information. Contact information for one institution's clinical trials office (Duke Cancer Center) could not be obtained from the publicly facing website; therefore they were not recruited to participate. Twenty-one (21) institutions explicitly denied to participate for various reasons, including upcoming NCI reporting deadlines, lack of time and/or resources to assist with data abstraction, and non-interest in the project subject. Twenty-three (23) institutions failed to respond to any contact and were determined to be non-responsive after three recruitment emails.

The geographic distribution of centers recruited for participation is illustrated in **Figure 4-2**. No centers from the northeastern United States participated in this project. The largest proportion of centers were from the Midwest (10 centers, or 62.5%). One center was from the southeast and five centers were from the west. Eleven (11) participating centers are comprehensive cancer centers. This represents 27% of all NCI-designated comprehensive cancer centers. Three participating centers are clinical centers, representing 15% of NCI-designated clinical centers.

Figure 4-2 All Centers with Participating Centers Highlighted

Participating centers indicated by black stars. Non-participating, NCI-designated cancer centers indicated by a gray dot.



NCI Cancer Center Support Grant (CCSG) aggregate information from designated cancer centers is publically available at <http://cancercenters.cancer.gov/DT/DT4>. As of calendar year 2013 (the most recent year available), all 61 NCI-designated centers had accrued a total of 42,897 subjects on interventional treatment trials. Of these, 7,472 were accrued by the 20 clinical centers (average 374 subjects per center) and 35,425 were accrued by the 41 comprehensive cancer centers (average 864 subjects per center). The range of enrolled subjects at clinical centers was 161-756 (median 305) subjects. The range of enrolled subjects at comprehensive cancer centers was 84-6,675 (median 512) subjects. On median, the number of subjects enrolled on an interventional agent/device trial in calendar year 2013 was 2.7 subjects. As the median time a trial accrues (as determined by this project) is 1.7 years (average 2.2 years), then the average trial accrues approximately 4.6-5.9 subjects (median number of subjects per trial multiplied by the median number of years accruing to median number of subjects per trial multiplied by the average number of years accruing). Expanding this to the first and third quartiles, the approximate range of subjects accrued to trial increases to 2.6-7.8 subjects. This project found the overall average accrual per study to be 8.5 subjects (median 4.0 subjects) with the range of average accrual at sites to be 2.0- 14.7 subjects.

As per the NIH Reporter website (<https://projectreporter.nih.gov/reporter.cfm>), the average amount of NCI funding awarded to clinical centers in calendar year 2013 was \$22.25 million (U.S.); median \$20.8 million (range \$9.4M-\$56.8M). For comprehensive cancer centers, the average was \$51.343 million (U.S.); median \$37.5 million (range \$11.8M-\$250.7M). To

maintain the anonymity of sites, detailed information regarding each site's NCI funding is not provided. The range of funding for these 16 institutions is \$4.1M-\$129.7M (median \$33.1M).

The number of interventional treatment or supportive care cancer clinical trials in the ClinicalTrials.gov website with completion dates after January 1, 2009, and start dates before April 1, 2015, is 20,933. The number of studies contributed for this project is 2,763, or 13% of the number of clinical trials ongoing during the study period. These 20,933 studies had a target enrollment of 2,909,785 subjects. The 49,319 subjects enrolled by participating institutions represent 1.7% of the total number of subjects from all studies. Considering only the studies contributed as part of this project, subject enrollment was 6.9% of the total national enrollment goals.

4.2.2 Disease Team Summary Statistics

On average, disease teams predicted accrual of 19 subjects per protocol with a median of 10 subjects per protocol. Both the average and the median per protocol disease team-predicted values were approximately double the actual values. Overall, disease teams predicted accrual 221% higher than actual accrual (109,076 subjects predicted versus 49,319 actually accrued; **Table 4-3**). The total disease team predicted accrual for all 16 institutions was above actual accrual (range 144%-535%). The median percent difference between the disease team predicted accrual to actual accrual was 273% (range 178%-4,650%; zero accruing protocols had a percent difference set to 25,000%)

Table 4-3: Disease Team Prediction Summary Statistics

	TOTAL ACCRUAL	AVERAGE ACCRUAL (MEDIAN)	TOTAL DISEASE TEAM ACCRUAL PREDICTION (% OF ACTUAL)	AVERAGE DISEASE TEAM ACCRUAL PREDICTION (MEDIAN)	MEDIAN % DIFFERENCE: DISEASE TEAM TO ACTUAL ACCRUAL*
ALL SITES	49319	8.5 ± 17.5 (4)	109076 (221%)	19.1 ± 44.3 (10)	273%
SITE 104647	3497	9.6 ± 16.0 (5)	7066 (202%)	19.4 ± 56.6 (10)	178%
SITE 173472	3407	8.4 ± 15.5 (3)	5886 (173%)	14.5 ± 31.6 (8)	200%
SITE 448155	6716	9.6 ± 14.2 (4)	9657 (144%)	13.9 ± 16.3 (10)	167%
SITE 494048	4029	7.3 ± 14.9 (3)	8694 (216%)	16.1 ± 21.0 (10)	300%
SITE 512786	6228	12.9 ± 26.5 (7)	12696 (204%)	26.3 ± 38.4 (18)	231%
SITE 560623	363	2.0 ± 3.5 (1)	1206 (332%)	9.3 ± 9.9 (5)	4650%
SITE 575415	3549	8.9 ± 14.4 (4)	6094 (172%)	14.9 ± 15.8 (10)	215%
SITE 598430	1338	4.1 ± 6.9 (2)	7156 (535%)	22.0 ± 68.2 (12)	683%
SITE 602591	2113	5.3 ± 7.5 (3)	6299 (295%)	15.9 ± 24.1 (10)	400%
SITE 689326	6705	14.7 ± 26.7 (7)	18834 (281%)	41.3 ± 107.0 (20)	267%
SITE 696337	831	4.6 ± 5.6 (2.5)	1917 (231%)	10.5 ± 6.6 (10)	333%
SITE 714415	5530	10.5 ± 26.7 (4)	8629 (156%)	16.3 ± 29.0 (10)	200%

SITE 715532	1313	5.8 ± 10.8 (2)	4756 (362%)	21.1 ± 33.8 (10)	600%
SITE 846594	1854	6.3 ± 9.7 (2)	6586 (355%)	23.9 ± 35.3 (14)	500%
SITE 997056	754	8.1 ± 13.4 (4)	1490 (198%)	16.0 ± 14.3 (11)	250%
SITE 998666	1092	5.7 ± 8.9 (2)	2110 (193%)	11.0 ± 11.7 (10)	300%

**Fold difference set at 250 for zero-accurring protocols, thus medians used to represent the difference between team prediction and actual prediction.*

4.2.3 Protocol Summary Statistics

Protocol summary statistics are listed in **Table 4-4**.

Table 4-4: Protocol Summary Statistics by Accrual Status

VARIABLE	ALL STUDIES	ZERO-ACCRUING PROTOCOLS	ACCRUING PROTOCOLS	P-VALUE
FREQUENCY (% OF ALL STUDIES)	5787	1053	4734	
# UNIQUE STUDIES	2763 (47.7%)			
TOTAL ACCRUAL	49319	0	49319	
RANDOMIZED DESIGN				<0.001
YES	2731 (47.2%)	544 (51.7%)	2187 (46.2%)	
NO	3056 (52.8%)	509 (48.3%)	2547 (53.8%)	
PLACEBO/OBSERVATION ARM				0.34
YES	628 (10.9%)	123 (11.7%)	505 (10.7%)	
NO	5159 (89.1%)	930 (88.3%)	4229 (89.3%)	
PHASE				<0.001
PILOT	15 (0.3%)	2 (0.2%)	13 (0.3%)	
PHASE 0	10 (0.2%)	0 (0.0%)	10 (0.2%)	
PHASE I	854 (14.8%)	109 (10.4%)	745 (15.7%)	
PHASE I/II	491 (8.5%)	78 (7.4%)	413 (8.7%)	
PHASE II	2419 (41.8%)	474 (45.0%)	1945 (41.1%)	

PHASE II/III	65 (1.1%)	20 (1.9%)	45 (1.0%)	
PHASE III	1765 (30.5%)	344 (32.7%)	1421 (30.0%)	
PHASE IV	34 (0.6%)	5 (0.5%)	29 (0.6%)	
NONE	134 (2.3%)	21 (2.0%)	113 (2.4%)	
OUTCOME/ENDPOINT				<0.001
SAFETY ONLY	677 (11.7%)	80 (7.6%)	597 (12.6%)	
EFFICACY	5030 (86.9%)	956 (90.2%)	4074 (86.1%)	
OTHER	80 (1.4%)	17 (1.6%)	63 (1.3%)	
PEDIATRIC SUBJECTS				<0.001
YES	554 (9.8%)	155 (14.7%)	339 (32.2%)	
NO	5233 (90.4%)	898 (85.3%)	4335 (91.6%)	
SPONSOR TYPE				<0.001
EXTERNALLY PEER REVIEWED	442 (7.6%)	36 (3.4%)	406 (8.6%)	
INSTITUTIONAL	644 (11.1%)	55 (5.2%)	589 (12.4%)	
INDUSTRY	2674 (46.2%)	451 (42.8%)	2223 (47.0%)	
NATIONAL GROUP	2027 (35.0%)	511 (48.5%)	1516 (32.0%)	
# NATIONAL SITES				<0.001
1-9	1695 (29.3%)	179 (17.0%)	1516 (32.0%)	
10-49	1524 (26.3%)	317 (30.1%)	1207 (25.5%)	
50-199	1468 (25.4%)	358 (34.0%)	1110 (23.4%)	
200+	925 (16.0%)	169 (16.0%)	756 (16.0%)	
UNKNOWN	175 (3.0%)	30 (2.8%)	145 (3.1%)	

USE OF LOCAL IRB				0.995
EXTERNAL	786 (13.4%)	143 (13.6%)	643 (13.6%)	
LOCAL	4999 (86.4%)	909 (86.3%)	4090 (86.4%)	
PRECISION MEDICINE STUDIES				0.01
YES	185 (3.2%)	46 (4.6%)	139 (3.2%)	
NO	5047 (87.2%)	955 (94.8%)	4092 (95.3%)	
CONDITIONAL	71 (1.2%)	6 (0.6%)	65 (1.5%)	
UNKNOWN*	484 (8.4%)	46	438	
PRIMARY PURPOSE				0.25
TREATMENT	5186 (89.6%)	980 (97.3%)	4206 (97.9%)	
SUPPORTIVE CARE	117 (2.0%)	27 (2.7%)	90 (2.1%)	
UNKNOWN*	484 (8.4%)	46	438	
AVERAGE ± SD (MEDIAN)				
NATIONAL ENROLLMENT GOAL	391.8 ± 1567.7 (104)	289.4 ± 684.4 (108)	414.5 ± 1705.7 (104)	<0.001**
TOTAL # MONTHS ACCRUING	41.8 ± 26.6 (35.1)	40.8 ± 28.5 (33.0)	42.0 ± 26.1 (35.7)	0.01**
# MONTHS ACCRUAL COMPLETED BEFORE LOCAL ENROLLMENT BEGINS	9.2 ± 11.0 (5.9)	12.8 ± 13.4 (8.8)	8.4 ± 10.3 (5.3)	<0.001**
DISEASE TEAM PREDICTION (MEDIAN)	19.1 ± 44.3 (10)	11.9 ± 36.0 (6)	20.7 ± 45.8 (12)	<0.001
ACTIVATION TIME, DAYS (N=3291)	190.1 ± 164.0 (155)	206.6 ± 179.0 (164)	186.5 ± 160.3 (150)	0.04**

AVERAGE # SUBJECTS PER SITE***	11.5 ± 30.1 (4.0)	8.9 ± 39.4 (2.4)	12.1 ± 27.6 (4.5)	0.02
--------------------------------	-------------------	------------------	-------------------	------

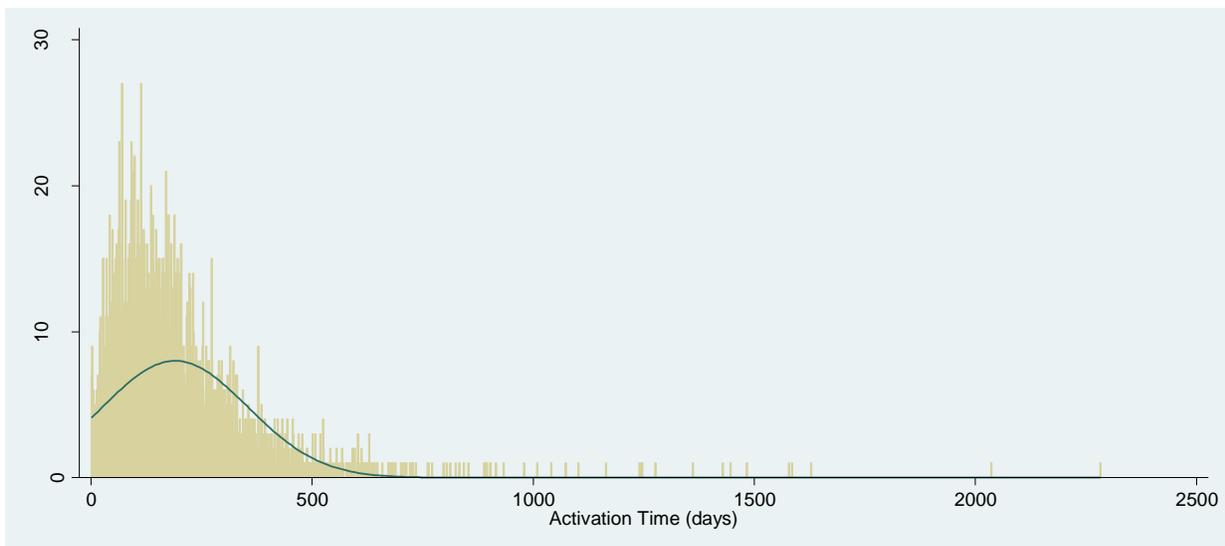
*Statistical significance between studies that accrued and studies that did not accrue using chi-square test (categorical variables) and two-sided t-test with unequal variances (continuous variables). Percentages are within accrual category (column). *Not included in chi-square analyses. **T-test utilized transformed variables. ***Only values under 1000 were assessed.*

Of the 5,787 protocols contributed, 2,763 (47.7%) were unique protocols and just over 3,000 being conducted at more than one participating institution. Of these 5,787 protocols, approximately 47% had a randomized design and 11% had a placebo and/or observation arm. Approximately 88% of protocols were either phase I, II, or phase III protocols with only 1.1% of protocols being phase 0, phase IV, or pilot studies. Eighty-seven percent (87%) of studies have an efficacy endpoint and just under 10% of studies involved pediatric subjects at the national level. Approximately 19% of studies were institutional-based, either sponsored by the institution or externally peer reviewed. Almost half, 46.2%, of studies were industry-sponsored, and 35% of studies were NCI National Group (cooperative group) sponsored. The largest category of sites was 1-9 sites at 29%, with 52% of studies split between 10-49 and 50-199 sites, 16% conducted at over 200 sites, and 3% of studies having an unknown number of national sites. The majority of protocols (86.4%) were overseen by the institutional IRB. The majority of protocols, approximately 90%, were treatment protocols, intended to treat a disease or condition, and a minority of protocols (4.4%) had a precision medicine component.

On average, national accrual to a study lasted for about three and one-half years, with a median of three years (35 months). On average, studies were open nationally for 9.2 months prior to opening locally (median 5.9 months). At the site level, this average ranged from 7.2-13.9 months (**Table 4-2**; page 130). On average, it took sites 190 days, or 6.3 months, to activate a trial (median 154 days). Activation time was right-skewed, with 129 studies (3.9%) activating in over 500 days (**Figure 4-3**). Histograms of activation time for each site are presented in

Appendix E-1. Once open studies, on average, accrued for 792 days, or 2.2 years (median 627 days, or 1.7 years; institutional average range: 609-998 days).

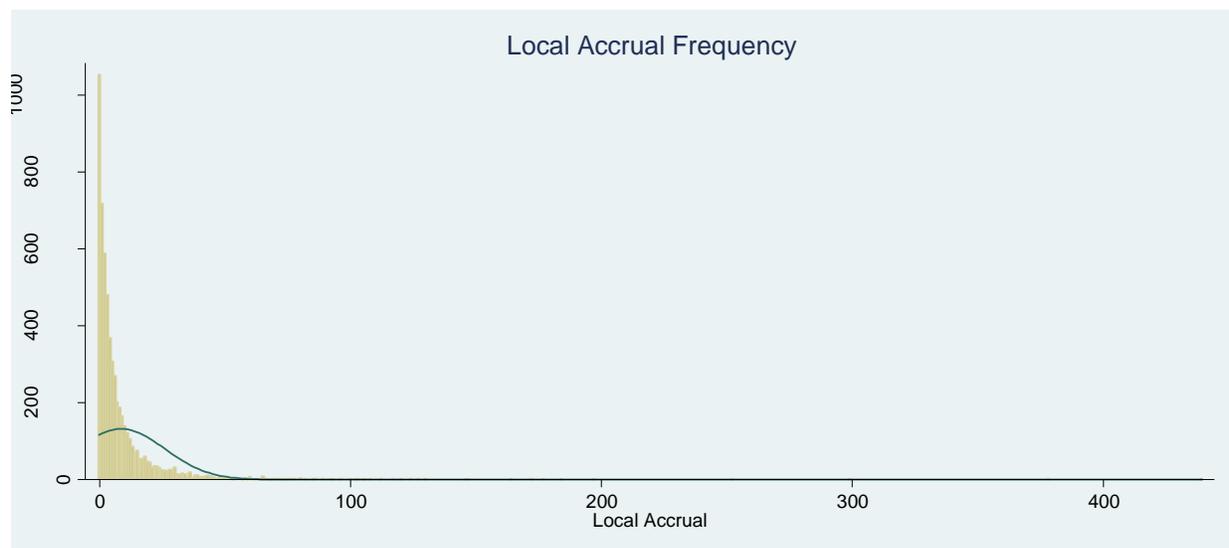
Figure 4-3: Distribution of Activation Time – All Sites



4.2.3.1 Zero-Accruing Protocol Summary Statistics

Clinical trial accrual was highly right-skewed (**Figure 4-4**).

Figure 4-4: Distribution of Actual Accrual – All Sites



Of the 5,787 protocols, 1,053, or 18.3%, did not accrue any subjects at the site level. The percentage of studies that accrued one subject was 12.4%. Conversely, 25% of studies accrued at least 10 subjects (range 10-439). Clinical trial accrual histograms for each site are presented in **Appendix E-2**. Zero-accruing studies were statistically more likely to be smaller studies (mean national enrollment goal of 289 subjects versus 414 for accruing studies; $p < 0.001$); however, the median enrollment for these studies was similar (104 versus 108 subjects, non-accruing versus accruing). The proportion of randomized trials to non-randomized trials was statistically higher in zero-accruing trials ($p < 0.001$; **Table 4-4**; page 138). Additionally, the distribution of trial phase was significantly different between zero-accruing trials and accruing

trials, with a higher proportion of zero-accruing protocols being phase II, II/III, and III studies. All phase I studies in this sample accrued subjects. A higher proportion of non-accruing studies had an efficacy endpoint ($p < 0.001$). Pediatric studies were more likely to not accrue subjects (14.7% of non-accruing studies were pediatric versus 7.3% of accruing studies). Zero-accruing studies were less likely to involve 1-9 sites (17% versus 32.0%) and more likely to involve between 10-49 or 50-199 sites (30.1% versus 25.5% and 34.0% versus 23.4%, respectively). The average number of subjects per site was statistically lower in zero-accruing studies (8.9 versus 12.1; $p = 0.02$). There was no difference between zero-accruing studies and accruing studies with the variables of placebo/observation arm, primary purpose (treatment versus supportive care), or the IRB of record. Accruing studies were nationally open to accrual longer than non-accruing studies (42 months versus 40.8 months). Protocols that did not accrue took statistically longer to activate (207 versus 187 days) and were nationally open to accrual longer (13 months versus 8 months) than accruing studies. Finally, disease teams predicted higher subject accrual, on average and on median for studies that accrued than those that did not accrue.

4.3 General Statistical Considerations

Continuous variables were reviewed to assess whether they fit a normal distribution. **Figure 3-1** (page 115) contains the distribution matrices for the continuous variables of months of national accrual completed prior to local opening, national enrollment goal, total months of planned national accrual, and the number of days to activate a study (activation time). None of the continuous variables were normally distributed. To best fit the variables to a normal distribution,

a square root transformation was used on the months of accrual completed and activation time. Log transformations were used on national enrollment goal and total months of national accrual.

4.4 Specific Aim One

4.4.1 Contributing Protocol Summary Statistics

Eleven sites submitted information regarding the use of a feasibility committee (68.75% of all sites; **Table 4-5**). Of these 11 sites, three had a 5-stage/3-gate process by utilizing a feasibility committee throughout the study (494048, 598430, and 714145). Two sites (104647 and 696337) instated a feasibility committee during the study period. However, no studies for 104647 utilizing the feasibility committee were submitted for this analysis due to the late instatement of the committee. The remaining six institutions (173472, 448155, 512786, 560623, 602591, and 998666) did not have feasibility committees and, therefore, utilized a 3-stage/2-gate process.

Of the total 5,787 protocols in this study, stage-gate classification was provided for 4,309 protocols, or 75% of studies. Of these 4,309 protocols, 34.4%, or 1,484 protocols, utilized a 5-stage/3-gate process which incorporated a feasibility committee as part of the study activation process (**Table 4-6**).

Table 4-5: Feasibility Committee, Regulatory Staffing Model, and Study Activation Information

	FEASIBILITY COMMITTEE	REGULATORY STAFF MODEL	ACTIVATION TIME (DAYS)	
			AVERAGE	MEDIAN
SITE 104647	2014-2015: National Group	Combined with other CRC/budget	Not provided	Not provided
SITE 173472	No	Not provided	279.2 ± 232.8	217
SITE 448155	No	Dedicated regulatory staff	160.7 ± 135.2	138
SITE 494048	Yes	Dedicated regulatory staff	124.0 ± 107.9	100
SITE 512786	No	Dedicated regulatory staff	227.4 ± 191.5	181.5
SITE 560623	No	Dedicated regulatory staff	167.8 ± 130.1	127.5
SITE 575415	Not provided	Not provided	Not provided	Not provided
SITE 598430	Yes	Not provided	Not provided	Not provided
SITE 602591	No	Dedicated regulatory staff	220.1 ± 131.9	202
SITE 689326	Not provided	Not provided	Not provided	Not provided
SITE 696337	Instated July 2013	Progressed to dedicated regulatory staff	155.6 ± 114.2	137
SITE 714415	Yes	Dedicated regulatory staff	184.4 ± 151.3	146
SITE 715532	Not provided	Not provided	Not provided	Not provided
SITE 846594	Not provided	Not provided	Not provided	Not provided
SITE 997056	Not provided	Not provided	Not provided	Not provided
SITE 998666	No	Dedicated regulatory staff	Not provided	Not provided

Table 4-6: Summary Statistics by Stage-Gate Classification

VARIABLE	ALL STUDIES	3-STAGE/2-GATE PROCESS	5-STAGE/3-GATE PROCESS	P-VALUE
# OF CONTRIBUTING PROTOCOLS	4309 (74.5% of all)	2825 (65.6%)	1484 (34.4%)	
TOTAL ACCRUAL	35144	23897 (68.0%)	11247 (32.0%)	
ZERO-ACCRUING PROTOCOLS				0.023
YES	828 (19.2%)	515 (18.2%)	313 (21.1%)	
NO	3481 (80.8%)	2310 (81.8%)	1171 (78.9%)	
RANDOMIZED DESIGN				0.009
YES	2066 (47.9%)	1314 (46.5%)	752 (50.7%)	
NO	2243 (52.1%)	1511 (53.5%)	732 (49.3%)	
PLACEBO/OBSERVATION ARM				0.455
YES	486 (11.2%)	326 (11.5%)	160 (10.8%)	
NO	3823 (88.7%)	2499 (88.5%)	1324 (89.2%)	
PHASE				<0.001
PILOT	15 (0.3%)	15 (0.5%)	0 (0.0%)	
PHASE 0	8 (0.2%)	7 (0.2%)	1 (6.7%)	
PHASE I	626 (14.5%)	437 (15.5%)	189 (12.7%)	
PHASE I/II	357 (8.3%)	264 (9.3%)	93 (6.3%)	
PHASE II	1850 (42.9%)	1224 (43.3%)	626 (42.2%)	

PHASE II/III	47 (1.1%)	29 (1.0%)	18 (1.2%)	
PHASE III	1277 (29.6%)	781 (27.6%)	496 (33.4%)	
PHASE IV	29 (0.7%)	14 (0.5%)	15 (1.0%)	
NONE	100 (2.3%)	54 (1.9%)	46 (3.1%)	
OUTCOME/ENDPOINT				0.07
SAFETY ONLY	507 (11.8%)	355 (12.6%)	152 (10.2%)	
EFFICACY	3735 (86.7%)	2428 (85.9%)	1307 (88.1%)	
OTHER	67 (1.6%)	42 (1.5%)	25 (1.7%)	
PEDIATRIC SUBJECTS				<0.001
YES	342 (7.9%)	166 (5.9%)	176 (11.9%)	
NO	3947 (91.6%)	2659 (94.1%)	1308 (88.1%)	
SPONSOR TYPE				<0.001
EXTERNALLY PEER REVIEWED	301 (7.0%)	232 (8.2%)	69 (4.6%)	
INSTITUTIONAL	481 (11.2%)	299 (10.9%)	182 (12.3%)	
INDUSTRY	2114 (49.1%)	1440 (51.0%)	674 (45.4%)	
NATIONAL GROUP	1413 (32.8%)	854 (30.2%)	559 (37.7%)	
# NATIONAL SITES				<0.001
1-9	1234 (28.6%)	855 (30.3%)	379 (25.5%)	
10-49	1165 (27.0%)	754 (26.7%)	411 (27.7%)	
50-199	1098 (25.5%)	700 (24.8%)	398 (26.8%)	
200+	683 (15.9%)	417 (14.8%)	266 (17.9%)	
UNKNOWN	129 (3.0%)	99 (3.5%)	30 (2.0%)	

USE OF LOCAL IRB				<0.001
EXTERNAL	729 (16.9%)	585 (20.7%)	144 (9.7%)	
LOCAL	3580 (83.1%)	2240 (79.3%)	1340 (90.3%)	
PRECISION MEDICINE STUDIES				0.07
YES	185 (3.2%)	82 (2.9%)	69 (4.6%)	
NO	5047 (87.2%)	2226 (78.8%)	1402 (94.5%)	
CONDITIONAL	71 (1.1%)	33 (1.2%)	13 (0.9%)	
UNKNOWN*	484 (8.4%)	484	0	
PRIMARY PURPOSE				0.12
TREATMENT	5186	2293 (81.2%)	1442 (97.2%)	
SUPPORTIVE CARE	117	48 (1.7%)	42 (2.8%)	
UNKNOWN*	484	484	0	
AVERAGE ± SD (MEDIAN)				
NATIONAL ENROLLMENT GOAL	391.8 ± 1567.7 (104)	345.1 ± 798.65 (100)	450.1 ± 2717.4 (115)	0.04**
ACCRUAL	8.2 ± 17.1 (4)	8.5 ± 16.1 (4)	7.6 ± 18.8 (3)	0.13
TOTAL # MONTHS ACCRUING	41.6 ± 26.3 (34.1)	40.7 ± 25.5 (34.5)	41.8 ± 27.7 (34.0)	0.51**
# MONTHS ACCRUAL COMPLETED BEFORE LOCAL ENROLLMENT BEGINS	9.2 ± 11.1 (5.9)	9.4 ± 10.8 (6.3)	8.9 ± 11.5 (5.3)	0.07**
DISEASE TEAM PREDICTION	16.8 ± 34.4 (10)	16.6 ± 31.8 (10)	17.2 ± 38.7 (10)	0.69

ACTIVATION TIME, DAYS	190.1 ± 164.0 (154)	209.6 ± 174.7 (174)	152.1 ± 133.0 (114)	<0.001**
-----------------------	------------------------	---------------------	---------------------	----------

*Statistical significance between stage-gate processes using chi-square test (categorical variables) and two-sided t-test with unequal variances (continuous variables). Percentages are within stage-gate category (column). *Not included in chi-square analyses. **T-test conducted on transformed variables.*

These studies accrued 11,247 of the total 35,144 subjects (32.0%). There was a statistical difference in zero-accruing protocols by stage-gate classification, with the 5-stage/3-gate process (feasibility committee) having a higher proportion of zero-accruing studies (21.1% versus 18.2%; $p=0.02$). There was no statistical difference in whether the protocol utilized a placebo/observation arm, was a precision medicine protocol, was treatment or supportive care (primary purpose), or what the outcome/endpoint classification was. Studies with a feasibility committee were more likely to be randomized (50.7% versus 46.5%; $p=0.009$), were more likely to be phase III and less likely to be phase I or I/II ($p<0.001$), more likely to be pediatric studies (11.9% versus 5.9%; $p<0.001$), more likely to be institutionally sponsored or national group sponsored protocols ($p<0.001$), have a larger number of national sites ($p<0.001$), and utilize a local IRB (90.3% versus 79.3%; $p<0.001$).

By t-test, there was no statistical difference in the average accrual per protocol between protocols approved through the 5-stage/3-gate process versus those with a 3-stage/2-gate process, nor was there a difference in the average or median number of subjects that the disease team predicted per protocol. The number of months that protocols accrued nationally and the number of months the protocol was accruing nationally prior to site activation did not differ between the two groups. The number of days to activate a protocol was significantly shorter in the 5-stage/3-gate process than the 3-stage/2-gate process (152 versus 210 days, respectively).

4.4.2 Univariate Analyses and Reduced Models

Univariate analyses found that local clinical trial accrual and activation time were associated significantly with the stage-gate classification ($p < 0.001$; **Table 4-7**). The likelihood that a study would accrue zero subjects or at least four subjects was not statistically different between stage-gate classifications (0.64 and 0.92, respectively).

Table 4-7: Adjusted, Univariate Models Assessing the Impact of the 5-stage/3-gate versus the 3-stage/2-gate Framework

Table 4-7a: Stage-Gate Process Impact on Clinical Trial Accrual

```

Mixed-effects nbinomial regression      Number of obs      =      4,309
Overdispersion:      mean
Group variable:      siteid             Number of groups   =      11

                                           Obs per group:
                                           min =      179
                                           avg =     391.7
                                           max =      697

Integration method: mvaghermite         Integration pts.   =      7

Log likelihood = -13237.505             Wald chi2(1)      =      15.75
                                           Prob > chi2       =      0.0001
-----
local_accrual |      Coef.   Std. Err.   z    P>|z|   [95% Conf. Interval]
-----+-----
feasibility |
  Yes       | -0.1789546  .045092   -3.97  0.000   -0.2673332  -0.0905759
  _cons     |  2.230763   .0278573  80.08  0.000   2.176164    2.285363
-----+-----
  /lnalpha  |  .5611938   .0232311  24.16  0.000   .5156617    .6067259
-----+-----
siteid      |
  var(_cons)|  .9707698   .8037396                .1915896    4.918815
-----+-----
LR test versus nbinomial model:  chibar2(01) = 0.00      Prob >= chibar2 = 1.0000

Akaike's information criterion and Bayesian information criterion
-----+-----
Model |      Obs   ll(null)  ll(model)   df      AIC      BIC
-----+-----
. |      4,309          .  -13237.5    4    26483.01  26508.48
-----+-----

```

Table 4-7b: Stage-Gate Process Impact on the Likelihood of Zero-Accruing Studies

```

Mixed-effects logistic regression          Number of obs   =    4,309
Group variable:          siteid           Number of groups =    11

                                           Obs per group:
                                           min =    179
                                           avg =   391.7
                                           max =    697

Integration method: mvaghermite           Integration pts. =    7
                                           Wald chi2(1)    =    0.22
                                           Prob > chi2     =    0.6385
Log likelihood = -2047.0931
-----
      nonaccr |          Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
feasibility |
  Yes       |   .1226852   .2611094     0.47  0.638    - .38908   .6344503
  _cons    |  -1.40666   .1839444    -7.65  0.000    -1.767184 -1.046135
-----+-----
siteid      |
  var(_cons)|   .278137   .1278213                .113   .6846035
-----+-----
LR test versus logistic model: chibar2(01) = 117.82      Prob >= chibar2 = 0.0000

Akaike's information criterion and Bayesian information criterion

```

```

-----+-----
      Model |          Obs   ll(null)   ll(model)      df          AIC          BIC
-----+-----
      .    |          4,309          .   -2047.093         3    4100.186    4119.291
-----+-----

```

Table 4-7c: Stage-Gate Process Impact on the Likelihood of Accruing at Least Four Subjects

```

Mixed-effects logistic regression          Number of obs   =    4,309
Group variable:          siteid           Number of groups =    11

                                           Obs per group:
                                           min =    179
                                           avg =   391.7
                                           max =    697

Integration method: mvaghermite           Integration pts. =    7
                                           Wald chi2(1)    =    0.01
                                           Prob > chi2     =    0.9243
Log likelihood = -2891.2671
-----
actualaccrual_binary |          Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
feasibility |
  Yes       |  -.0229072   .2409405    -0.10  0.924    -.4951419   .4493275
  _cons    |  -.1426572   .191091    -0.75  0.455    -.5171885   .2318742
-----+-----

```

```

siteid          |
      var(_cons)|   .3239486   .1490363                .1314835   .798144
-----

```

LR test versus logistic model: chibar2(01) = 178.42 Prob >= chibar2 = 0.0000

Akaike's information criterion and Bayesian information criterion

```

-----
Model |      Obs  ll(null)  ll(model)    df      AIC      BIC
-----+-----
. |      4,309          . -2891.267      3    5788.534    5807.64
-----

```

Table 4-7d: Stage-Gate Process Impact on Activation Timeline

```

Mixed-effects ML regression          Number of obs   =      3,291
Group variable: siteid              Number of groups =           8

Obs per group:
    min =      100
    avg =     411.4
    max =      696

Wald chi2(1) =      13.65
Prob > chi2  =      0.0002

Log likelihood = -9839.9609

```

```

-----
activationtime_trans |      Coef.  Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
feasibility |
  Yes | -2.814833   .7619804   -3.69   0.000   -4.308287   -1.321379
  _cons |  13.79284   .5403907   25.52   0.000    12.7337    14.85199
-----

```

```

-----
Random-effects Parameters |      Estimate  Std. Err.    [95% Conf. Interval]
-----+-----
siteid: Identity
  var(_cons) |      1.715872   .9035741     .6112852    4.816434
-----+-----
  var(Residual) |      22.96498   .5668301    21.88045    24.10325
-----

```

LR test versus linear model: chibar2(01) = 231.68 Prob >= chibar2 = 0.0000

Akaike's information criterion and Bayesian information criterion

```

-----
Model |      Obs  ll(null)  ll(model)    df      AIC      BIC
-----+-----
. |      3,291          . -9839.961      4    19687.92    19712.32
-----

```

Reduced models focusing on the two independent variables associated with the stage-gate classification, accrual and activation time, were run to determine effect modifiers and interactions. As activation time was not normally distributed, the variable was square root transformed prior to regression analysis. Independent variables statistically associated with clinical trial accrual and stage-gate process include study phase, IRB of record, the number of national sites, and the number of months opened to enrollment nationally prior to local activation (Table 4-8).

Table 4-8: Variable Association with Clinical Trial Accrual by Stage-Gate Process

Variable	STAGE-GATE PROCESS
Study phase	Effect Modifier
IRB of record	Effect Modifier
Sponsor type	Not an Effect Modifier
Pediatric subjects	Not an Effect Modifier
Randomized	Not an Effect Modifier
Placebo	Not an Effect Modifier
Primary endpoint	Not an Effect Modifier
National enrollment goal¥	Not an Effect Modifier
# Sites nationally	Effect Modifier
Total # months accruing¥	Not an Effect Modifier
# Months open prior to local open¥	Effect Modifier

Adjusted, reduced negative binomial regression models of the associated variable with stage-gate process and outcome. Change in stage-gate beta coefficient $\geq 10\%$ were defined as “Effect Modifier.”

¥ Transformed variable used (log transformation for national enrollment and total number of months accruing; square root transformation for number of months of national accrual completed).

In addition, national enrollment goal was an effect modifier when combined with the number of months accruing prior to local activation. National enrollment goal was a statistically significant interaction with the number of national sites and sponsor type (**Table 4-9**). No individual independent variable was associated with activation time (**Table 4-10**). However, national enrollment goal paired with number of national sites, months of accrual completed prior to local activation, total number of months accruing, and sponsor type had statistically significant interactions with activation time.

Table 4-9: Effect Modification and Interaction Analyses by Stage-Gate Process

VARIABLE INTERACTION PAIR WITH NATIONAL ENROLLMENT	# SITES	MONTHS NATIONAL ACCRUAL COMPLETED	TOTAL MONTHS PLANNED ACCRUAL	SPONSOR TYPE
CLINICAL TRIAL ACCRUAL	Interaction	Effect Modifier	Neither	Interaction
ACTIVATION TIME	Interaction	Interaction	Interaction	Interaction

Tests with an interaction term that had a corresponding p-value < 0.05 defined as “Interaction.” If no interaction, models with a change in the beta coefficient of stage-gate process (feasibility committee) $\geq 10\%$ were defined as “Effect Modifier,” otherwise listed as “Neither.”

Table 4-10: Variable Association with Activation Time* by Stage-Gate Process

Variable	STAGE-GATE PROCESS
Study phase	Not an Effect Modifier
IRB of record	Not an Effect Modifier
Sponsor type	Not an Effect Modifier
Pediatric subjects	Not an Effect Modifier
Randomized	Not an Effect Modifier
Placebo	Not an Effect Modifier
Primary endpoint	Not an Effect Modifier
National enrollment goal‡	Not an Effect Modifier
# Sites nationally	Not an Effect Modifier
Total # months accruing‡	Not an Effect Modifier
# Months open prior to local open‡	Not an Effect Modifier

Adjusted, reduced linear regression models of the associated variable with stage-gate process and outcome. Change in stage-gate beta coefficient $\geq 10\%$ were defined as “Effect Modifier.”

**The dependent variable, activation time, is square root transformed to fit the observations to a normal distribution.*

‡ Transformed variable used (log transformation for national enrollment and total number of months accruing; square root transformation for number of months of national accrual completed).

4.4.3 Multivariate Analyses

To assess the impact of stage-gate process on clinical trial accrual, an adjusted negative binomial regression model was run with institution as the random effect. The best fit model, as indicated

with the lowest combination of AIC, BIC, and LR values, included all effect modifiers found in univariate analyses along with national enrollment goal and the interaction terms of national enrollment goal/number of national sites and national enrollment goal/sponsor type. The overall model was statistically significant ($p < 0.001$; **Table 4-11**); however the random effects model was not significantly different than the unadjusted negative binomial model, according to a likelihood ratio test ($p = 1.00$).

Table 4-11: Full Model Assessing the Impact of Stage-Gate Process on Clinical Trial Accrual

```

Mixed-effects nbinomial regression      Number of obs      =      4,309
Overdispersion:      mean
Group variable:      siteid             Number of groups   =      11

Obs per group:
      min =      179
      avg =     391.7
      max =      697

Integration method: mvaghermite         Integration pts.   =      7

Wald chi2(26)      =     2261.96
Prob > chi2       =      0.0000
Log likelihood = -12204.555
  
```

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
feasibility						
Yes	-.2777754	.0368729	-7.53	0.000	-.3500449	-.2055058
phase						
Phase 0	.2551013	.4492091	0.57	0.570	-.6253323	1.135535
Phase 1	.3256238	.2741846	1.19	0.235	-.2117683	.8630158
Phase 1/2	.2836178	.2766942	1.03	0.305	-.2586928	.8259285
Phase 2	.0772046	.2723487	0.28	0.777	-.456589	.6109981
Phase 2/3	-.062967	.3145285	-0.20	0.841	-.6794315	.5534976
Phase 3	.0258935	.2778582	0.09	0.926	-.5186985	.5704855
Phase 4	.2662529	.338369	0.79	0.431	-.3969382	.9294439
None	.3086327	.2910449	1.06	0.289	-.2618049	.8790702
irb						
Local	.1338625	.0489625	2.73	0.006	.0378977	.2298272
sponsor_type						
Instituitonal	-.4606043	.2577046	-1.79	0.074	-.9656961	.0444874
Industry	-.4945821	.2372453	-2.08	0.037	-.9595743	-.02959
National Group	-1.719353	.2704264	-6.36	0.000	-2.249379	-1.189327
natenroll_trans	.6580189	.0556724	11.82	0.000	.548903	.7671349
natsitescat						
10-49 sites	-.2114541	.2231416	-0.95	0.343	-.6488036	.2258954
50-199 sites	.0935645	.277132	0.34	0.736	-.4496042	.6367332

200+ sites		1.229288	.3226304	3.81	0.000	.5969446	1.861632
Unknown		.0172598	.3544754	0.05	0.961	-.6774992	.7120187
moaccrdone_trans		-.2556715	.0114414	-22.35	0.000	-.2780962	-.2332467
natsitescat#c.natenroll_trans							
10-49 sites		-.0971948	.0528922	-1.84	0.066	-.2008616	.0064721
50-199 sites		-.2314509	.0580235	-3.99	0.000	-.3451749	-.1177268
200+ sites		-.4272427	.0608705	-7.02	0.000	-.5465467	-.3079388
Unknown		-.1579042	.0744893	-2.12	0.034	-.3039007	-.0119078
sponsor_type#c.natenroll_trans							
Instititutional		.1690689	.0675149	2.50	0.012	.0367422	.3013957
Industry		-.0812207	.0584953	-1.39	0.165	-.1958693	.0334279
National Group		.1620102	.0623818	2.60	0.009	.0397441	.2842764
_cons		.5881068	.3484196	1.69	0.091	-.094783	1.270997
/lnalpha		-.0241734	.0275557	-0.88	0.380	-.0781815	.0298347
siteid							
var(_cons)		.5285417	.4396658			.1035165	2.698663

LR test versus nbinomial model: $\text{chibar2}(01) = 0.00$ Prob >= $\text{chibar2} = 1.0000$

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	4,309	.	-12204.55	29	24467.11	24651.79

A mixed linear regression model was used with institution as the random effect. No independent variables were associated with activation time in univariate analyses. Interaction models showed significant interactions by national enrollment goal/number of national sites, national enrollment goal/months of national enrollment completed prior to local activation, national enrollment goal/total months of national enrollment planned, and national enrollment goal/sponsor type (**Table 4-9**). The best fit, full model included two of these interactions: national enrollment goal/months of national enrollment completed prior to local activation and national enrollment goal/sponsor type ($p < 0.001$; **Table 4-12**).

Table 4-9: Effect Modification and Interaction Analyses by Stage-Gate Process

VARIABLE INTERACTION PAIR WITH NATIONAL ENROLLMENT	# SITES	MONTHS NATIONAL ACCRUAL COMPLETED	TOTAL MONTHS PLANNED ACCRUAL	SPONSOR TYPE
CLINICAL TRIAL ACCRUAL	Interaction	Effect Modifier	Neither	Interaction
ACTIVATION TIME	Interaction	Interaction	Interaction	Interaction

Tests with an interaction term that had a corresponding p-value < 0.05 defined as “Interaction.” If no interaction, models with a change in the beta coefficient of stage-gate process (feasibility committee) $\geq 10\%$ were defined as “Effect Modifier,” otherwise listed as “Neither.”

Table 4-12: Full Model Assessing the Impact of Stage-Gate Process on Activation Time

Mixed-effects ML regression
 Group variable: siteid
 Number of obs = 3,291
 Number of groups = 8
 Obs per group:
 min = 100
 avg = 411.4
 max = 696
 Wald chi2(14) = 997.65
 Prob > chi2 = 0.0000
 Log likelihood = -9410.8477

activationtime_trans	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
feasibility						
Yes	-3.126758	.6903422	-4.53	0.000	-4.479804	-1.773712
natenroll_trans	-1.331208	.3007388	-4.43	0.000	-1.920645	-.741771
moaccrdone_trans	.1252523	.1705441	0.73	0.463	-.209008	.4595126
natsitescat						
10-49 sites	-.2853882	.2263421	-1.26	0.207	-.7290105	.1582341
50-199 sites	-.147319	.2779505	-0.53	0.596	-.692092	.3974541
200+ sites	.315389	.3559312	0.89	0.376	-.3822232	1.013001
Unknown	-.9454167	.4650712	-2.03	0.042	-1.856939	-.0338939
sponsor_type						
Instituitonal	.6964668	1.321792	0.53	0.598	-1.894199	3.287132
Industry	-3.766294	1.223059	-3.08	0.002	-6.163445	-1.369142
National Group	-6.969231	1.263021	-5.52	0.000	-9.444706	-4.493756
natenroll_trans	0	(omitted)				
moaccrdone_trans	0	(omitted)				
c.natenroll_trans#c.moaccrdone_trans	.1860665	.0352477	5.28	0.000	.1169822	.2551507
natenroll_trans	0	(omitted)				
sponsor_type#c.natenroll_trans						
Instituitonal	.3176154	.3419432	0.93	0.353	-.3525811	.9878118
Industry	.903499	.3011569	3.00	0.003	.3132424	1.493756
National Group	.7842738	.3041336	2.58	0.010	.1881829	1.380365
_cons	17.93924	1.295313	13.85	0.000	15.40047	20.47801

```

-----
Random-effects Parameters | Estimate Std. Err. [95% Conf. Interval]
-----+-----
siteid: Identity         |
var(_cons)              | 1.501917 .8069441 .5239809 4.30503
-----+-----
var(Residual)           | 17.68811 .4366123 16.85273 18.56489
-----

```

LR test versus linear model: $\chi^2(01) = 223.42$ Prob $\geq \chi^2 = 0.0000$

Akaike's information criterion and Bayesian information criterion

```

-----
Model | Obs ll(null) ll(model) df AIC BIC
-----+-----
. | 3,291 . -9410.848 17 18855.7 18959.38
-----

```

When controlling for effect modifiers and interactions, stage-gate process continues to have a strong, significant effect on clinical trial accrual ($z = -7.53$; $p < 0.001$). Clinical trial accrual goes up an absolute value of 0.76 (after exponentiating the beta-coefficient of -0.278) with the utilization of a feasibility committee (5-stage/3-gate process) holding all other variables constant. Stage-gate process also has a strong, significant effect on activation time ($z = -4.53$; $p < 0.001$). Studies with a 5-stage/3-gate process have a reduced activation time by 3.12 days on the square root of activation time.

4.4.4 Site-Specific Analysis

One institution, site 696337, instituted a feasibility committee during the study period. To assess the impact of the feasibility committee (i.e. moving to a 5-stage/3-gate process from a 3-stage/2-gate process), the same regression methodology for all centers was used for this site. Forty-two percent (42%) of submitted protocols ($n=76$) were activated during the timeframe when a feasibility committee was in place.

Univariate analyses examining the relationship between the use of a feasibility committee and clinical trial accrual, whether a study is likely to accrue any subjects, or whether a study is likely to accrue at least four subjects showed no statistical association ($p=0.94$, $p=0.96$, $p=0.80$, respectively; **Table 4-13**).

Table 4-13: Single Site Analysis: 696337 Univariate Models Assessing the Impact of the 5-stage/3-gate versus the 3-stage/2-gate Framework

Table 4-13a: Stage-Gate Process Impact on Clinical Trial Accrual

Negative binomial regression	Number of obs	=	182
Dispersion = mean	LR chi2(1)	=	0.01
Log likelihood = -475.23533	Prob > chi2	=	0.9360
	Pseudo R2	=	0.0000

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
feasibility					
Yes	.0147716	.1839142	0.08	0.936	-.3456936 .3752369
_cons	1.512428	.1189212	12.72	0.000	1.279347 1.745509
/lnalpha	.2458468	.1308302			-.0105758 .5022694
alpha	1.278704	.1672931			.98948 1.652467

LR test of alpha=0: chibar2(01) = 559.00 Prob >= chibar2 = 0.000

Table 4-13b: Stage-Gate Process Impact on the Likelihood of Zero-Accruing Studies

Logistic regression	Number of obs	=	182
Log likelihood = -93.246639	LR chi2(1)	=	0.00
	Prob > chi2	=	0.9611
	Pseudo R2	=	0.0000

nonaccr	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
feasibility					
Yes	.0180185	.3694942	0.05	0.961	-.7061768 .7422139
_cons	-1.339774	.239498	-5.59	0.000	-1.809182 -.8703668

Table 4-13c: Stage-Gate Process Impact on the Likelihood of Accruing at Least Four Subjects

Logistic regression	Number of obs	=	182
Log likelihood = -123.95725	LR chi2(1)	=	0.07
	Prob > chi2	=	0.7969
	Pseudo R2	=	0.0003

actualaccrual_binary	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
feasibility					
Yes	.0782522	.3039986	0.26	0.797	-.5175741 .6740785
_cons	-.3429448	.19712	-1.74	0.082	-.7292929 .0434034

Table 4-13d: Stage-Gate Process Impact on Activation Timeline

Source	SS	df	MS	Number of obs	=	100
Model	217.07483	1	217.07483	F(1, 98)	=	9.53
Residual	2231.99822	98	22.775492	Prob > F	=	0.0026
Total	2449.07305	99	24.7381116	R-squared	=	0.0886
				Adj R-squared	=	0.0793
				Root MSE	=	4.7724

activation~s	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feasibility						
Yes	-3.402547	1.102131	-3.09	0.003	-5.58969	-1.215403
_cons	14.00133	.9544735	14.67	0.000	12.10721	15.89545

There was a statistically significant reduction in the amount of time it took to activate protocols that had feasibility committee review (p=0.003). Multivariate analyses showed that phase was the only independent variable that had a modifying effect on feasibility committee (**Table 4-14**). Thus the full model included stage-gate process and phase. After controlling for phase, the use of a feasibility committee remained a statistically significant factor in activation time, reducing the square root of activation time by 4.2 days (z =-3.89; p>0.001), holding all other variables constant (**Table 4-15**).

Table 4-14: Single Site: Variable Association with Activation Time* by Stage-Gate Process

Site 696337	STAGE-GATE PROCESS
Study phase	Effect Modifier
IRB of record	Collinear – single IRB used
Sponsor type	Not an Effect Modifier
Pediatric subjects	Not an Effect Modifier
Randomized	Not an Effect Modifier

Placebo	Not an Effect Modifier
Primary endpoint	Not an Effect Modifier
National enrollment goal¥	Not an Effect Modifier
# Sites nationally	Not an Effect Modifier
Total # months accruing¥	Not an Effect Modifier
# Months open prior to local open¥	Not an Effect Modifier

Reduced linear regression models of the associated variable with stage-gate process and outcome. Change in stage-gate beta coefficient $\geq 10\%$ were defined as “Effect Modifier.”

**The dependent variable, activation time, is square root transformed to fit the observations to a normal distribution.*

¥ Transformed variable used (log transformation for national enrollment and total number of months accruing; square root transformation for number of months of national accrual completed).

Table 4-15: Single Site Analysis: 696337 Full Model Assessing the Impact of Stage-Gate

Process on Activation Time

Source	SS	df	MS	Number of obs	=	100
Model	582.390822	6	97.065137	F(6, 93)	=	4.84
Residual	1866.68223	93	20.0718519	Prob > F	=	0.0002
				R-squared	=	0.2378
				Adj R-squared	=	0.1886
Total	2449.07305	99	24.7381116	Root MSE	=	4.4802

activation~s	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
feasibility					
Yes	-4.189684	1.078081	-3.89	0.000	-6.330539 -2.048828
phase					
Phase 1/2	2.408739	1.86388	1.29	0.199	-1.292558 6.110036
Phase 2	3.624257	1.226335	2.96	0.004	1.188998 6.059516
Phase 2/3	-3.985967	3.403633	-1.17	0.245	-10.74491 2.772974
Phase 3	4.215335	1.2701	3.32	0.001	1.693167 6.737502
None	-2.874203	4.595926	-0.63	0.533	-12.0008 6.252394
_cons	11.85972	1.244505	9.53	0.000	9.388379 14.33106

4.4.5 Specific Aim One Summary

- Hypothesis 1a Finding: **Reject the null hypothesis** ($p < 0.001$)
- Hypothesis 1b Finding: **Fail to reject the null hypothesis** ($p = 0.64$).

After controlling for other contributing factors and adjusting for institution, adding a feasibility committee and utilizing the 5-stage/3-gate process significantly increased subject accrual by an average of 0.8 subjects (Hypothesis 1a) and decreased study activation time. There was, however, no association between the use of a feasibility committee and the likelihood of whether a study would accrue zero subjects (Hypothesis 1b) or whether they would accrue at least four subjects. In a case study of the single institution that implemented a feasibility committee during the study period, the only statistically significant difference was in activation time, which was reduced at approximately the same magnitude as the overall model, after controlling for effect modifiers.

4.5 Specific Aim Two

4.5.1 Institutional Infrastructure

Nine centers (56.25%) provided infrastructure information regarding regulatory functions surrounding study activation (**Table 4-54**). Seven centers (43.75%) provided all the information needed to complete the second Specific Aim (centers 448155, 494048, 512786, 560623, 602591, 696337, and 714145; centers 104647 and 998666 did not provide the start dates for study activation). Three centers are comprehensive cancer centers, two are clinical centers, and two are not designated. Five institutions provided information regarding the date of study activation.

Four indicated that the start date for activation is officially listed based around PRMS timelines (three are upon PRMS submission; one is upon PRMS approval) and one institution indicated that the start date is PRMS feasibility committee approval, prior to PRMS scientific review. Allowance of concurrent PRMS and IRB reviews differs between institutions. Two institutions do not allow concurrent PRMS and IRB submission (448155, 998666). Two institutions allow concurrent submission under specific circumstances (494048 allows it if there is a grant deadline, 104647 allows concurrent submission but IRB cannot review until PRMS has approved). Two institutions allow concurrent submission (512786, 714145). One institution did not allow concurrent submission, but changed practice during the study period (January 2012) to allow it (696337). Clinical trial office utilization varied among and within institutions with varying levels of disease team utilization.

4.5.2 Contributing Protocols

The number of protocols submitted for this aim is 2,133 (36.9% of all submitted protocols) from seven institutions (sites 448155, 494048, 512786, 560623, 602591, 696337, and 714145). Protocols accrued a total of 14,229 subjects (28.9% of all studies). Forty-two (42) protocols, or 20.7% of the Specific Aim 2 protocols, accrued no subjects. Almost half (47.4%) accrued at least four subjects. Protocols had similar distributions of protocol characteristics compared to the overall study population (**Table 4-16**). The average activation time for contributing protocols was 176.9 days (standard deviation = 130.4 days; median 149 days). Average accrual per protocol was 6.7 ± 14.2 subjects (median: 3). Study accrual and activation timelines were slightly lower in this population than the average and median activation times for the entire study

population. The median disease team prediction and number of months of accrual completed prior to local enrollment were the same between the Specific Aim 2 population and overall study population. Due to the low number of institutions that submitted complete and consistent data, specific variables regarding centralization of staff were not generated.

Table 4-16: Summary Statistics for Institutions Participating in Specific Aim Two

VARIABLE	ALL STUDIES	SPECIFIC AIM 2 STUDIES
# STUDIES	5787	2133 (36.9%)
TOTAL ACCRUAL	49319	14229 (28.9%)
RANDOMIZED DESIGN		
YES	2731 (47.2%)	1017 (47.7%)
NO	3056 (52.8%)	1116 (52.3%)
PLACEBO/OBSERVATION ARM		
YES	628 (10.9%)	270 (12.7%)
NO	5159 (89.1%)	1863 (87.3%)
PHASE		
PILOT	15 (0.3%)	12 (0.6%)
PHASE 0	10 (0.2%)	4 (0.2%)
PHASE I	854 (14.8%)	345 (16.2%)
PHASE I/II	491 (8.5%)	193 (9.1%)
PHASE II	2419 (41.8%)	951 (44.6%)
PHASE II/III	65 (1.1%)	32 (1.5%)
PHASE III	1765 (30.5%)	543 (25.5%)
PHASE IV	34 (0.6%)	16 (0.8%)
NONE	134 (2.3%)	37 (1.7%)

OUTCOME/ENDPOINT		
SAFETY ONLY	677 (11.7%)	297 (13.9%)
EFFICACY	5030 (86.9%)	1810 (84.9%)
OTHER	80 (1.4%)	26 (1.2%)
PEDIATRIC SUBJECTS		
YES	554 (9.8%)	158 (7.4%)
NO	5233 (90.4%)	1975 (92.6%)
SPONSOR TYPE		
EXTERNALLY PEER REVIEWED	442 (7.6%)	93 (4.4%)
INSTITUTIONAL	644 (11.1%)	233 (10.9%)
INDUSTRY	2674 (46.2%)	1226 (57.5%)
NATIONAL GROUP	2027 (35.0%)	581 (27.2%)
# NATIONAL SITES		
1-9	1695 (29.3%)	579 (27.1%)
10-49	1524 (26.3%)	634 (29.7%)
50-199	1468 (25.4%)	551 (25.8%)
200+	925 (16.0%)	306 (14.4%)
UNKNOWN	175 (3.0%)	63 (3.0%)
USE OF LOCAL IRB		
EXTERNAL	786 (13.4%)	248 (11.6%)
LOCAL	4999 (86.4%)	1885 (88.4%)
PRECISION MEDICINE STUDIES		
YES	185 (3.2%)	95 (4.5%)
NO	5047 (87.2%)	1617 (75.8%)
CONDITIONAL	71 (1.2%)	25 (1.2%)
UNKNOWN	484 (8.4%)	396 (18.6%)
PRIMARY PURPOSE		
TREATMENT	5186 (89.6%)	1691 (79.3%)

SUPPORTIVE CARE	117 (2.0%)	46 (2.2%)
UNKNOWN	484 (8.4%)	396 (18.9%)
AVERAGE ± SD (MEDIAN)		
LOCAL ACCRUAL	8.5 ± 17.5 (4)	6.7 ± 14.2 (3)
NATIONAL ENROLLMENT GOAL	391.8 ± 1567.7 (104)	277 ± 564.3 (102)
TOTAL # MONTHS ACCRUING NATIONALLY	41.8 ± 26.6 (35.1)	36.7 ± 22.4 (32.0)
TOTAL # DAYS ACCRUING LOCALLY	791.6 ± 629.7 (627)	586 ± 415.5 (490)
# MONTHS ACCRUAL COMPLETED BEFORE LOCAL ENROLLMENT BEGINS	9.2 ± 11.0 (5.9)	9.3 ± 10.8 (5.9)
DISEASE TEAM PREDICTION (MEDIAN)	19.1 ± 44.3 (10)	15.1 ± 21.7 (10)
ACTIVATION TIME, DAYS (N=3291)	190.1 ± 164.0 (155)	176.9 ± 130.4 (149)

4.5.3 Protocol Workload

Protocol workload was categorized two ways (number of protocols per staff member and number of protocols per FTE) and distributions plotted in **Figures 4-5** and **4-6** (accrual) and **Figure 4-7** (activation time). The average number of protocols per staff member for all sites was 5.2±3.0 (median 4.7; **Table 4-17**). The average number of protocols per FTE for all sites was 5.3±3.0 (median 4.7). Site averages ranged from 3.2-10.1 protocols per staff member/FTE.

Table 4-17: Protocol Workload Summary Statistics by Site

	AVERAGE PROTOCOLS PER STAFF MEMBER (MEDIAN)	AVERAGE PROTOCOLS PER FULL-TIME EQUIVALENT (FTE) (MEDIAN)
ALL SITES	5.2±3.0 (4.7)	5.3±3.0 (4.7)
SITE 448155	3.4±0.8 (3.6)	3.4±0.8 (3.7)
SITE 494048	5.5±1.7 (5.3)	5.5±1.7 (5.3)
SITE 512786	3.2±0.9 (3.1)	3.2±0.9 (3.1)
SITE 560623	5.7±2.4 (5)	5.7±2.4 (5)
SITE 602591	4.7±1.5 (4.8)	5.2±1.8 (5.3)
SITE 696337	4.0±2.3 (4.8)	4.0±2.3 (4.8)
SITE 714415	10.1±3.7 (9.8)	10.1±3.7 (9.8)

Average number of protocols per staff member and average number of protocols per FTE are statistically different between sites, by ANOVA analysis ($p < 0.001$).

Distribution data for each site's outcomes (clinical trial accrual, zero-accruing protocols, and activation time) by protocol workload are presented in **Appendix F-1** through **Appendix F-7**.

While protocol workload (number of protocols per staff and number of protocols per FTE) are continuous variables, neither have a linear relationship when correlated with local accrual or activation time (all ρ values between -0.06 and 0.07; **Figures 4-8** and **4-9**, respectively).

Figure 4-5: Protocols per Full-Time Equivalent versus Overall Protocol Accrual

Figure 4-5a: All Sites

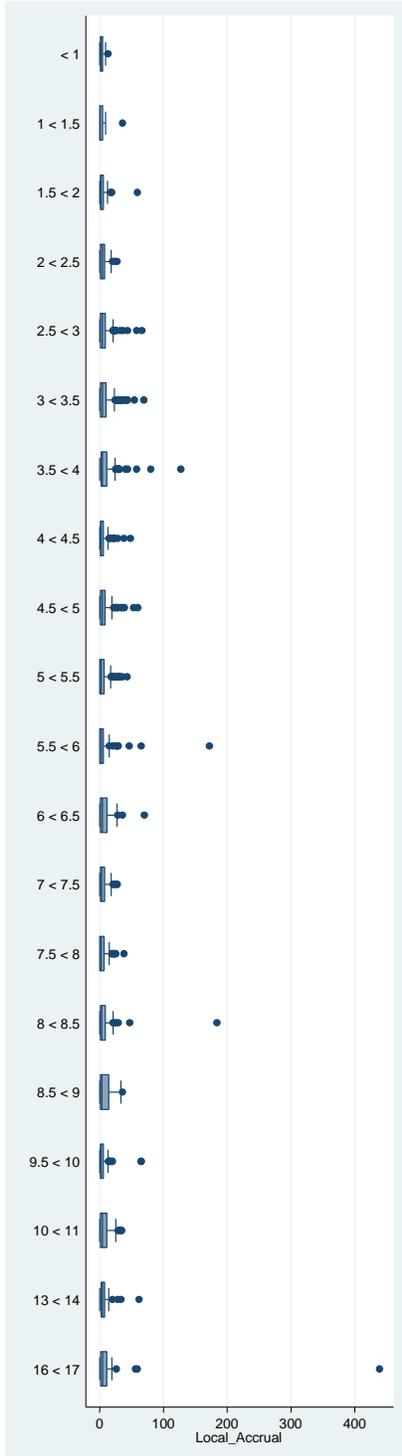


Figure 4-5b: All Sites: Accrual < 100

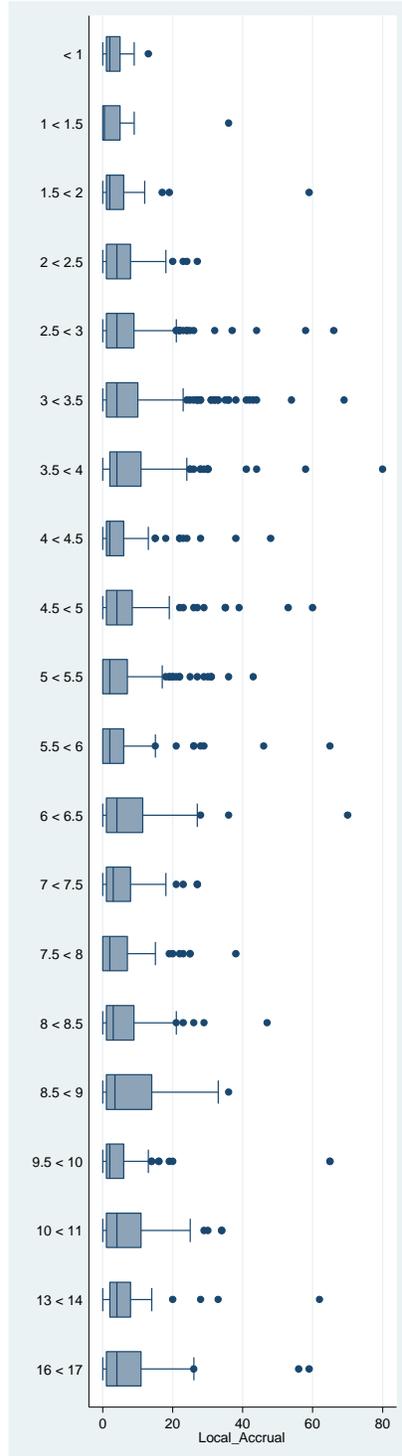


Figure 4-6: Protocols per Staff Member versus Overall Protocol Accrual – All Sites

Figure 4-6a: All Sites

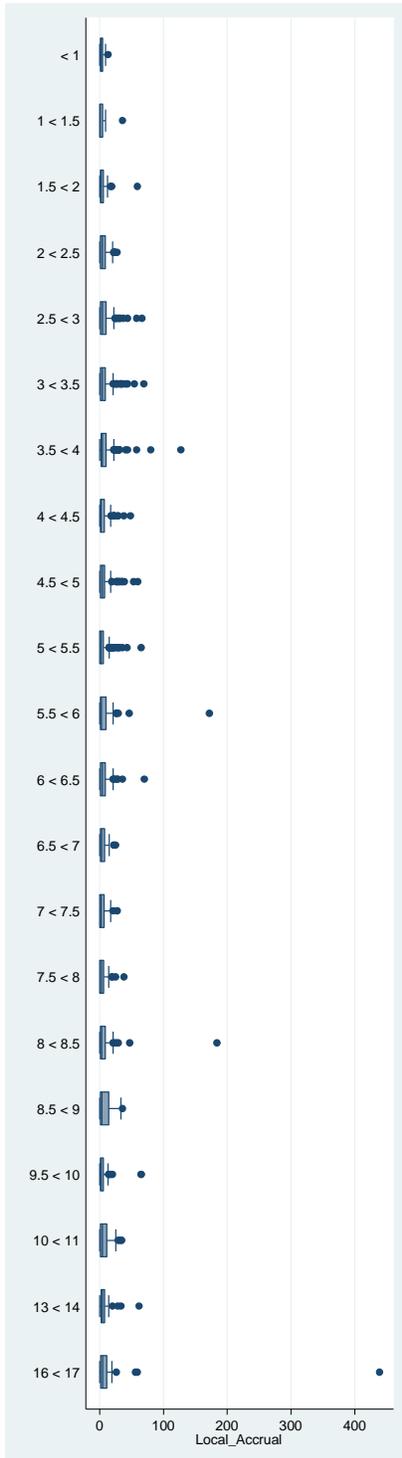


Figure 4-6b: All Sites: Accrual <100

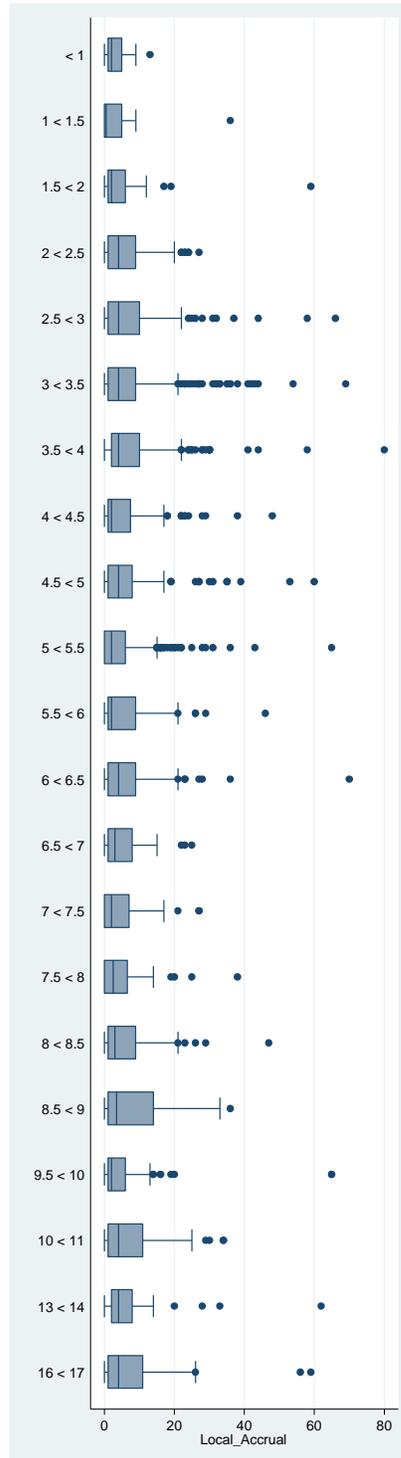


Figure 4-7: Protocol Workload versus Time to Study Activation – All Sites

Figure 4-7a: Protocols per FTE

Figure 4-7b: Protocols per Staff Member

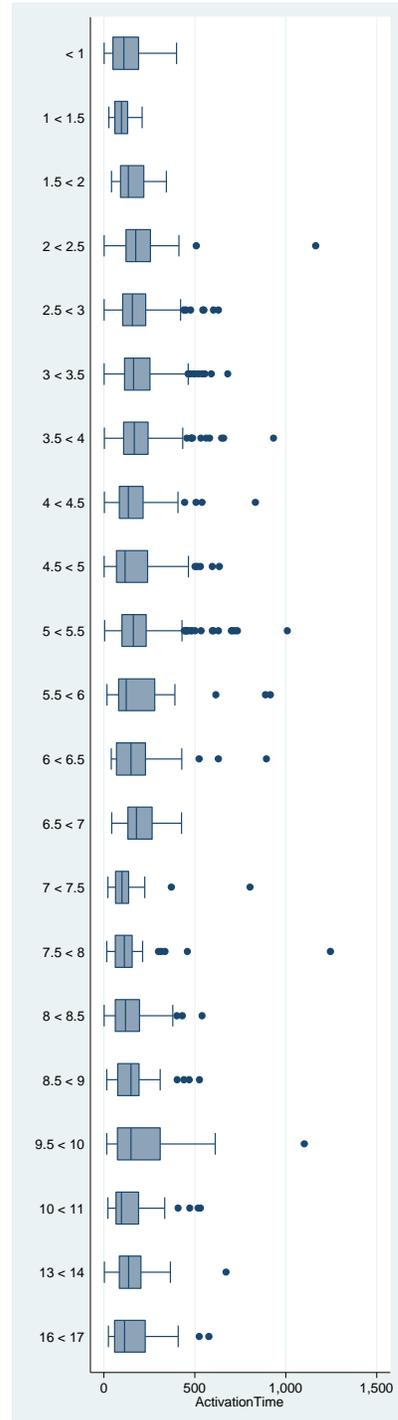
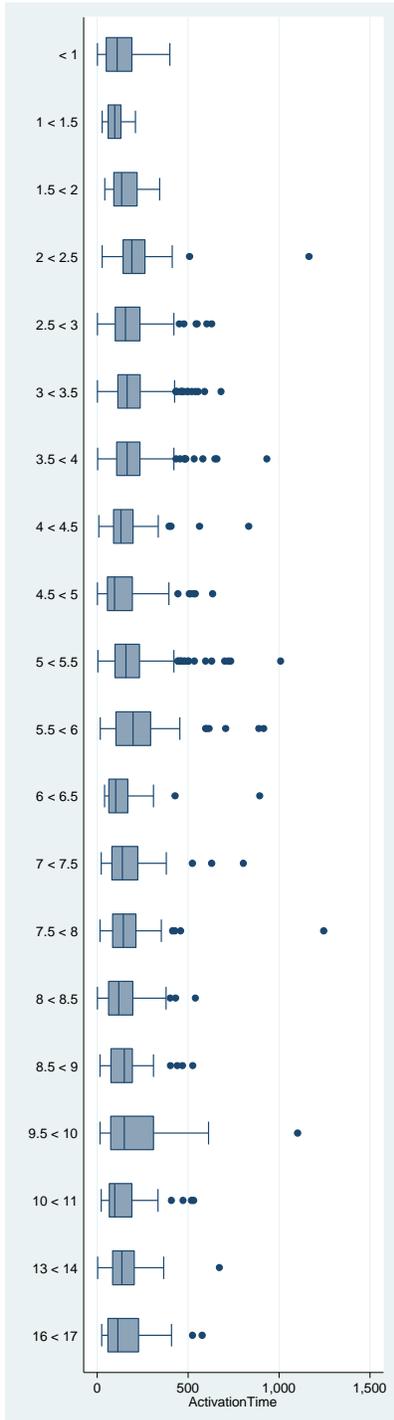


Figure 4-8 Mean Clinical Trial Accrual by Protocol Workload

Figure 4-8a: Protocols per Staff Member

Figure 4-8b: Protocols per FTE

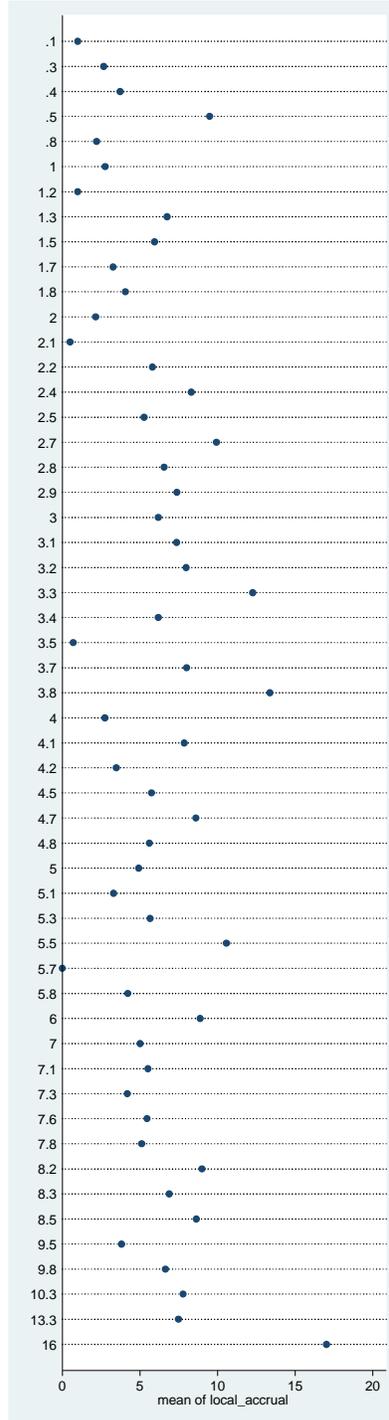
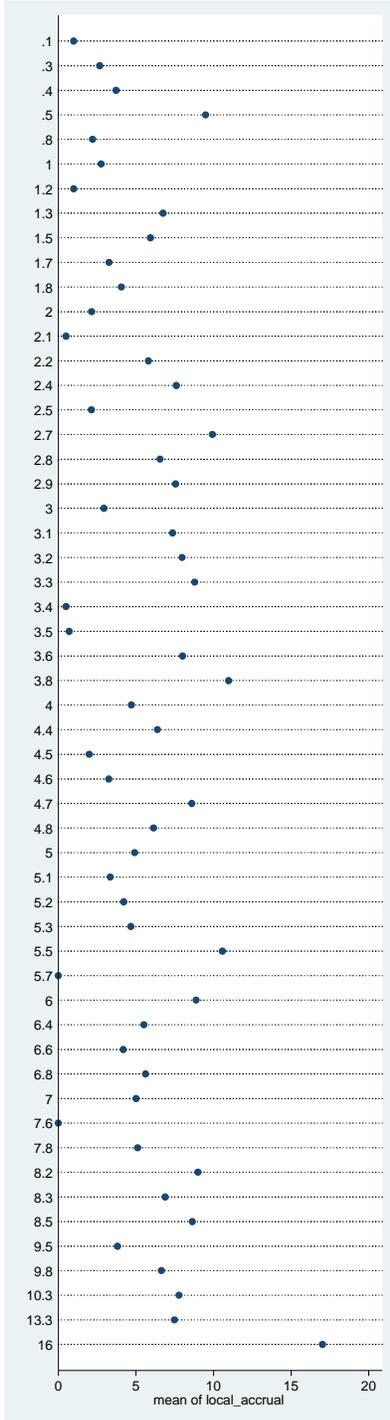
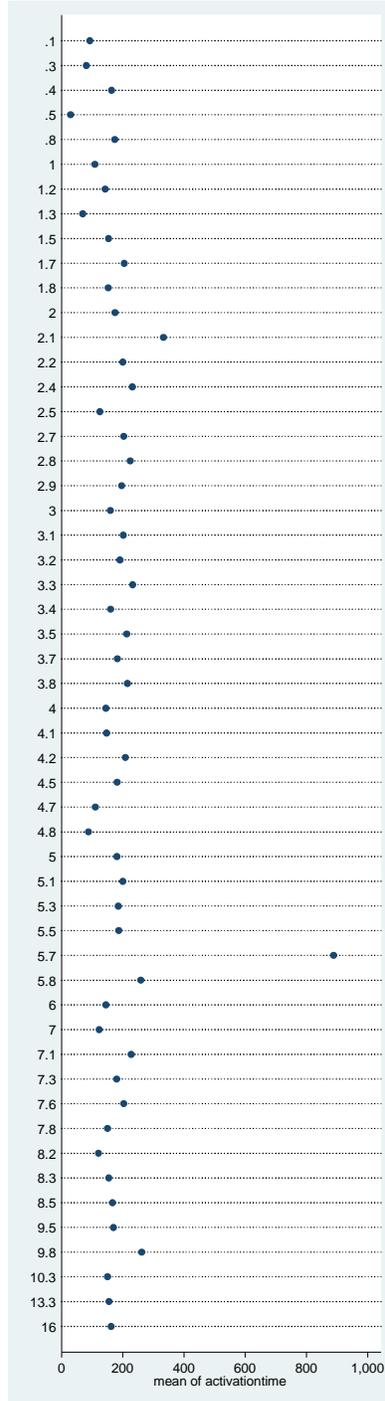
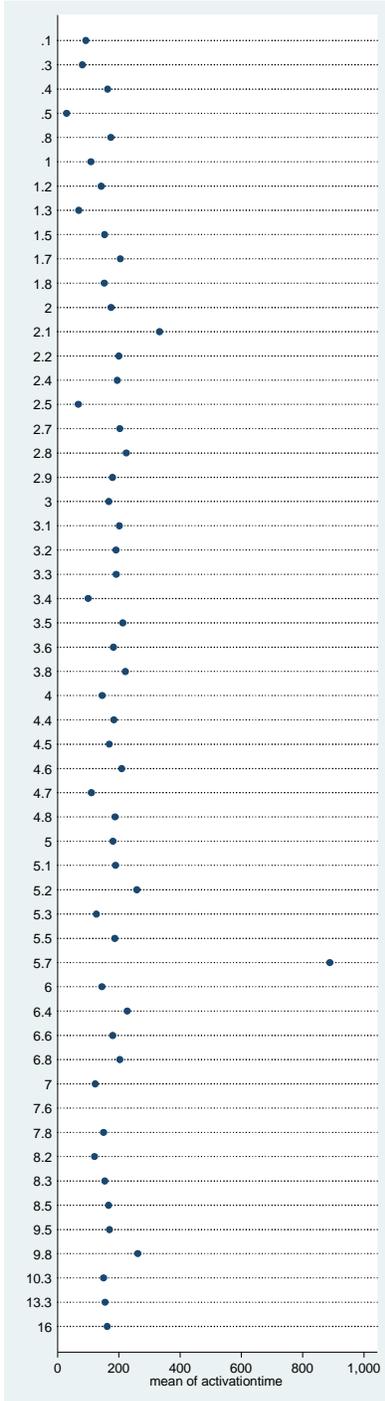


Figure 4-9 Mean Activation Time by Protocol Workload

Figure 4-9a: Protocols per Staff Member

Figure 4-9b: Protocols per FTE



Therefore, to assess the impact protocol workload has on these values, protocol workload was categorized as less than one protocol per staff member/FTE, in 0.5 increment categories for workloads between 1.0 and 10.0 protocols per staff member/FTE, and in 1.0 increment categories for workloads over 10.0 protocols per staff member/FTE. Two methods of protocol workload analysis were considered. The first was a study characteristic analysis that looked at the aggregate metric per workload category and chose the protocol workload corresponding to the most optimal outcome, regardless of any other protocol or institutional factors. The second analysis was regression modeling controlling for factors significantly and independently associated with the outcome.

Over all institutions, the number of protocols per staff member ranged from 0.11-16 (median 4.7; average 5.2 ± 3.0) and the number of protocols per FTE ranged from 0.8-16 (median 4.7; average 5.3 ± 3.0). The number of protocols per FTE was not statistically different than the number of protocols per staff member ($p=0.38$; t-test with equal variances) due to a number of institutions reporting dedicated regulatory personnel. Only one institution (714145) had workloads exceeding 10 protocols per staff member/FTE. Four outcomes defining “trial success” were considered: the frequency of zero-accruing protocols, the frequency of studies accruing at least four subjects, activation time, and clinical trial accrual. Summary results for all participating institutions are presented in tabular format in **Table 4-18** (protocols per staff member) and **Table 4-19** (protocols per FTE).

Table 4-18: Efficiency Outcomes by Protocol Workload (Protocols per Staff Member) – Participating Sites

# PROTOCOLS PER STAFF MEMBER	# ZERO-ACCRUING (%)	# 4+ ACCRUING (%)	AVERAGE TRIAL ACCRUAL (MEDIAN)	AVERAGE ACTIVATION TIME, DAYS (MEDIAN)
ALL STUDIES (N=2133)	42 (20.7%)	1012 (47.4%)	6.7 ± 14.2 (3)	176.9 ± 130.4 (149)
<1	7 (23.3%)	10 (33.3%)	3.1 ± 3.3 (2)	128.8 ± 97.4 (109)
1.0 < 1.5	8 (50.0%)	6 (37.5%)	4.3 ± 9.0 (0.5)	97.3 ± 50.7 (96)
1.5 < 2.0	16 (23.9%)	25 (37.3%)	4.4 ± 7.9 (2)	161 ± 79.0 (134)
2.0 < 2.5	22 (20.6%)	62 (57.9%)	6.3 ± 6.6 (4)	199.1 ± 136.6 (175)
2.5 < 3.0	33 (19.6%)	88 (52.4%)	7.2 ± 9.8 (4)	184.1 ± 117.1 (156.5)
3.0 < 3.5	51 (18.4%)	156 (56.1%)	7.6 ± 10.1 (4)	190.8 ± 114.8 (163.5)
3.5 < 4.0	31 (15.9%)	109 (55.9%)	8.4 ± 13.5 (4)	195.7 ± 132.1 (167)
4.0 < 4.5	38 (22.1%)	67 (39.0%)	5.2 ± 7.3 (2)	157.4 ± 114.1 (134.5)
4.5 < 5.0	32 (20.9%)	78 (51.0%)	6.4 ± 9.4 (4)	161.3 ± 129.6 (116)
5.0 < 5.5	80 (26.3%)	119 (39.1%)	4.7 ± 7.1 (2)	188.9 ± 134.1 (161)
5.5 < 6.0	10 (21.7%)	21 (45.7%)	10.3 ± 26.2 (2.5)	202.1 ± 194.9 (123)
6.0 < 6.5	14 (17.5%)	43 (53.8%)	7.6 ± 10.4 (4)	178.0 ± 140.9 (148.5)
6.5 < 7.0	10 (18.5%)	24 (44.4%)	5.1 ± 5.9 (3)	194.4 ± 94.0 (179)
7.0 < 7.5	11 (26.2%)	16 (38.1%)	5.0 ± 7.1 (2)	123.3 ± 125.1 (99.5)

7.5 < 8.0	15 (31.3%)	17 (35.4%)	5.0 ± 7.6 (2.5)	150.0 ± 184.8 (112)
8.0 < 8.5	12 (12.6%)	39 (41.1%)	8.0 ± 19.8 (3)	137.4 ± 107.2 (119)
8.5 < 9.0	4 (11.8%)	17 (50.0%)	8.6 ± 10.9 (3.5)	166.8 ± 130.2 (149)
9.0 < 9.5	No studies	No studies	No studies	No studies
9.5 < 10.0	28 (24.4%)	42 (36.5%)	4.8 ± 9.2 (2)	201.2 ± 170.3 (149)
10 < 11	7 (17.1%)	23 (56.1%)	7.8 ± 9.8 (4)	150.1 ± 132.1 (97)
11 < 12	No studies	No studies	No studies	No studies
12 < 13	No studies	No studies	No studies	No studies
13 < 14	5 (12.5%)	24 (60.0%)	7.5 ± 11.3 (4)	155.9 ± 115.9 (135.5)
14 < 15	No studies	No studies	No studies	No studies
15 < 16	No studies	No studies	No studies	No studies
16 < 17	8 (16.7%)	26 (54.2%)	17.0 ± 63.3 (4.5)	162.5 ± 134.0 (113)
NOT SPECIFIED (N=885)	115 (13.0%)	543 (61.4%)	13.1 ± 25.5 (6)	214.4 ± 210.4 (166)

Study characteristic (unadjusted) analysis, category of protocols per staff member (protocol workload) that best maximized or minimized the efficiency outcomes.

Table 4-19: Efficiency Outcomes by Protocol Workload (Protocols per FTE) – Participating Sites

# PROTOCOLS PER FTE	# ZERO-ACCRUING (%)	# 4+ ACCRUING (%)	AVERAGE TRIAL ACCRUAL (MEDIAN)	AVERAGE ACTIVATION TIME, DAYS (MEDIAN)
ALL STUDIES (N=2133)	442 (20.7%)	1012 (47.4%)	6.7 ± 14.2 (3)	176.9 ± 130.4 (149)
<1	7 (23.3%)	10 (33.3%)	3.1 ± 3.3 (2)	128.8 ± 97.4 (109)
1.0 < 1.5	8 (50.0%)	6 (37.5%)	4.3 ± 9.0 (0.5)	97.3 ± 50.7 (96)
1.5 < 2.0	16 (23.9%)	23 (37.3%)	4.4 ± 7.9 (2)	161.3 ± 79.0 (134)
2.0 < 2.5	17 (22.4%)	44 (57.9%)	6.0 ± 6.3 (4)	217.5 ± 144.5 (191)
2.5 < 3.0	36 (22.2%)	86 (53.1%)	7.0 ± 9.8 (4)	184.6 ± 118.6 (154.5)
3.0 < 3.5	51 (16.4%)	176 (56.6%)	7.7 ± 9.9 (4)	187.9 ± 113.7 (164)
3.5 < 4.0	29 (16.1%)	101 (56.1%)	8.7 ± 14.0 (4)	192.1 ± 131.3 (164.5)
4.0 < 4.5	34 (24.6%)	46 (33.3%)	4.6 ± 7.2 (2)	154.4 ± 110.8 (130.5)
4.5 < 5.0	22 (16.2%)	72 (52.9%)	6.9 ± 9.6 (4)	139.7 ± 125.6 (94.5)
5.0 < 5.5	81 (25.2%)	136 (42.2%)	4.9 ± 6.5 (2)	184.5 ± 128.1 (157.5)
5.5 < 6.0	27 (27.6%)	39 (39.8%)	7.1 ± 19.4 (2.5)	232.0 ± 170.3 (196.5)
6.0 < 6.5	10 (20.8%)	25 (52.1%)	8.9 ± 12.6 (4)	144.9 ± 139.9 (102)
6.5 < 7.0	No studies	No studies	No studies	No studies
7.0 < 7.5	19 (20.0%)	44 (46.3%)	5.0 ± 5.9 (3)	171.0 ± 125.4 (138)

7.5 < 8.0	21 (25.9%)	31 (38.3%)	5.3 ± 7.3 (2)	172.1 ± 157.8 (143)
8.0 < 8.5	12 (12.6%)	39 (41.1%)	8.0 ± 19.8 (3)	137.4 ± 107.2 (119)
8.5 < 9.0	4 (11.8%)	17 (50.0%)	8.6 ± 10.9 (3.5)	166.8 ± 130.2 (149)
9.0 < 9.5	No studies	No studies	No studies	No studies
9.5 < 10.0	28 (24.4%)	42 (36.5%)	4.8 ± 9.2 (2)	201.2 ± 170.3 (149)
10 < 11	7 (17.1%)	23 (56.1%)	7.8 ± 9.8 (4)	150.1 ± 132.1 (97)
11 < 12	No studies	No studies	No studies	No studies
12 < 13	No studies	No studies	No studies	No studies
13 < 14	5 (12.5%)	24 (60.0%)	7.5 ± 11.3 (4)	155.9 ± 115.9 (135.5)
14 < 15	No studies	No studies	No studies	No studies
15 < 16	No studies	No studies	No studies	No studies
16 < 17	9 (16.7%)	26 (54.2%)	17.0 ± 63.3 (4.5)	162.5 ± 134.0 (113)
NOT SPECIFIED (N=885)	115 (13.0%)	543 (61.4%)	13.1 ± 25.5 (6)	214.4 ± 210.4 (166)

Study characteristic (unadjusted) analysis, category of protocols per staff member (protocol workload) that best maximized or minimized the efficiency outcomes.

4.5.4 Regression Analyses

Univariate, adjusted regression analyses (using institution as the random effect) of protocol workload (number of protocols per staff member, categorical) showed statistically significant associations with clinical trial accrual, the likelihood that a study will accrue at least four subjects, and activation time ($p < 0.001$, $p = 0.002$, $p < 0.001$, respectively). There was no association between the number of protocols per staff member and likelihood that a study would accrue zero subjects ($p = 0.052$). In measuring protocol workload by FTE, the significant association between protocol workload remained with clinical trial accrual and activation time ($p < 0.001$), but not with the likelihood that a study would accrue at least four subjects ($p = 0.07$). Protocol workload measured in FTE had no relationship with the likelihood that a study would accrue zero subjects ($p = 0.13$). All univariate analyses are presented in **Table 4-20**.

Table 4-20: Univariate Adjusted Models Assessing Protocol Workload Impact – All Sites

Table 4-20a: Protocol Workload (Protocols per Staff Member) on Clinical Trial Accrual

Mixed-effects nbinomial regression	Number of obs	=	2,133
Overdispersion: mean			
Group variable: siteid	Number of groups	=	7
	Obs per group:		
	min	=	101
	avg	=	304.7
	max	=	477
Integration method: mvaghermite	Integration pts.	=	7
Log likelihood = -6173.6986	Wald chi2(20)	=	77.44
	Prob > chi2	=	0.0000

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
prot_per_staff_cat					
1 < 1.5	.500862	.4276851	1.17	0.242	-.3373854 1.339109
1.5 < 2	.3131881	.298365	1.05	0.294	-.2715965 .8979727
2 < 2.5	.5916385	.286345	2.07	0.039	.0304127 1.152864
2.5 < 3	.7427684	.2736032	2.71	0.007	.206516 1.279021
3 < 3.5	.7800366	.2684986	2.91	0.004	.2537891 1.306284
3.5 < 4	.8925157	.2731533	3.27	0.001	.357145 1.427886
4 < 4.5	.5531032	.2771221	2.00	0.046	.0099539 1.096252
4.5 < 5	.6761685	.2745039	2.46	0.014	.1381508 1.214186
5 < 5.5	.3769186	.2664865	1.41	0.157	-.1453853 .8992225
5.5 < 6	1.083379	.3188337	3.40	0.001	.458476 1.708281
6 < 6.5	.8365091	.2953784	2.83	0.005	.2575779 1.41544
6.5 < 7	.3707461	.3134404	1.18	0.237	-.2435858 .9850781
7 < 7.5	.3607918	.3278434	1.10	0.271	-.2817694 1.003353
7.5 < 8	.3568792	.3198029	1.12	0.264	-.269923 .9836813
8 < 8.5	.863007	.2855355	3.02	0.003	.3033677 1.422646
8.5 < 9	.9038228	.338776	2.67	0.008	.239834 1.567812
9.5 < 10	.6742193	.2907264	2.32	0.020	.104406 1.244033
10 < 11	.7982214	.3267581	2.44	0.015	.1577874 1.438655
13 < 14	.7581703	.3285112	2.31	0.021	.1143001 1.40204
16 < 17	1.581041	.3152515	5.02	0.000	.9631594 2.198922
_cons	1.253399	.2570644	4.88	0.000	.7495624 1.757236
/lnalpha	.4315909	.0349541	12.35	0.000	.363082 .5000997
siteid					
var(_cons)	.8008622	.8195542			.1077679 5.951498

LR test versus nbinomial model: chibar2(01) = 0.00	Prob >= chibar2 = 1.0000
--	--------------------------

Table 4-20c: Protocol Workload (Protocols per Staff Member) on the Likelihood of Accruing at

Least Four Subjects

```
Mixed-effects logistic regression
Group variable:      siteid
Number of obs       =      2,133
Number of groups    =          7

Obs per group:
    min =          101
    avg =         304.7
    max =          477

Integration method: mvaghermite
Integration pts.    =          7

Log likelihood = -1403.3072
Wald chi2(20)      =         42.24
Prob > chi2        =         0.0026
```

actualaccrual_binary	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

prot_per_staff_cat						
1 < 1.5	.9434005	.6708848	1.41	0.160	-.3715097	2.258311
1.5 < 2	.1004802	.4717767	0.21	0.831	-.8241852	1.025146
2 < 2.5	.8323681	.4477921	1.86	0.063	-.0452883	1.710024
2.5 < 3	.5728753	.4280569	1.34	0.181	-.2661009	1.411851
3 < 3.5	.761794	.420364	1.81	0.070	-.0621044	1.585692
3.5 < 4	.8590936	.4263832	2.01	0.044	.0233979	1.694789
4 < 4.5	.4025516	.4315686	0.93	0.351	-.4433073	1.248411
4.5 < 5	1.336051	.4382257	3.05	0.002	.4771448	2.194958
5 < 5.5	.6491991	.4212824	1.54	0.123	-.1764992	1.474897
5.5 < 6	1.020161	.5078656	2.01	0.045	.0247626	2.015559
6 < 6.5	1.566928	.4698496	3.33	0.001	.6460394	2.487816
6.5 < 7	1.196224	.502164	2.38	0.017	.2120006	2.180447
7 < 7.5	.6212864	.5253497	1.18	0.237	-.4083801	1.650953
7.5 < 8	.5014461	.5149359	0.97	0.330	-.5078098	1.510702
8 < 8.5	.6323262	.4601048	1.37	0.169	-.2694626	1.534115
8.5 < 9	.8659664	.5680286	1.52	0.127	-.2473492	1.979282
9.5 < 10	.7203027	.4840884	1.49	0.137	-.2284931	1.669099
10 < 11	1.111847	.5513477	2.02	0.044	.0312252	2.192468
13 < 14	1.272675	.5559929	2.29	0.022	.1829485	2.362401
16 < 17	1.033538	.5374206	1.92	0.054	-.0197867	2.086863
_cons	-1.036113	.4740297	-2.19	0.029	-1.965194	-.1070316

siteid						
var(_cons)	.4590463	.2672609			.146649	1.436924

LR test versus logistic model: chibar2(01) = 86.54 Prob >= chibar2 = 0.0000

Table 4-20d: Protocol Workload (Protocols per Staff Member) on Activation Timeline

```
Mixed-effects ML regression
Group variable: siteid
Number of obs = 2,132
Number of groups = 7
Obs per group:
    min = 100
    avg = 304.6
    max = 477
Wald chi2(20) = 57.83
Prob > chi2 = 0.0000
Log likelihood = -6145.5245
```

activationtime_trans	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
-----+-----						
prot_per_staff_cat						
1 < 1.5	-.1259547	1.348252	-0.09	0.926	-2.76848	2.51657
1.5 < 2	1.657058	.9462461	1.75	0.080	-.1975505	3.511666
2 < 2.5	2.144032	.8995926	2.38	0.017	.3808625	3.907201
2.5 < 3	1.828885	.8594479	2.13	0.033	.1443976	3.513371
3 < 3.5	1.638649	.8413234	1.95	0.051	-.0103141	3.287613
3.5 < 4	2.57682	.8540525	3.02	0.003	.9029076	4.250732
4 < 4.5	2.194869	.8628842	2.54	0.011	.5036469	3.886091
4.5 < 5	1.342103	.8770442	1.53	0.126	-.3768718	3.061078
5 < 5.5	2.912709	.8383579	3.47	0.001	1.269558	4.55586
5.5 < 6	3.677495	1.035407	3.55	0.000	1.648134	5.706855
6 < 6.5	2.438918	.9439106	2.58	0.010	.5888872	4.288949
6.5 < 7	1.218229	1.019143	1.20	0.232	-.7792551	3.215712
7 < 7.5	2.238043	1.065335	2.10	0.036	.150025	4.32606
7.5 < 8	3.054885	1.043148	2.93	0.003	1.010352	5.099419
8 < 8.5	1.286359	.9263402	1.39	0.165	-.5292342	3.101953
8.5 < 9	.6874337	1.165655	0.59	0.555	-1.597208	2.972075
9.5 < 10	2.14965	.9509026	2.26	0.024	.2859152	4.013385
10 < 11	.0201322	1.125172	0.02	0.986	-2.185164	2.225428
13 < 14	.4570314	1.130165	0.40	0.686	-1.758051	2.672114
16 < 17	.4835802	1.095591	0.44	0.659	-1.663739	2.630899
_cons	10.54337	.9492317	11.11	0.000	8.682908	12.40383
-----+-----						

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
-----+-----				
siteid: Identity				
var(_cons)	1.921249	1.067428	.6466422	5.708254
-----+-----				
var(Residual)	18.46854	.5665858	17.39078	19.61309
-----+-----				
LR test versus linear model: chibar2(01) = 163.09	Prob >= chibar2 = 0.0000			

Table 4-20e: Protocol Workload (Protocols per FTE) on Clinical Trial Accrual

Mixed-effects nbinomial regression	Number of obs	=	2,133
Overdispersion: mean			
Group variable: siteid	Number of groups	=	7
	Obs per group:		
	min	=	101
	avg	=	304.7
	max	=	477
Integration method: mvaghermite	Integration pts.	=	7
Log likelihood = -6173.3476	Wald chi2(19)	=	78.02
	Prob > chi2	=	0.0000

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
prot_per_fte_cat					
1 < 1.5	.4923023	.4278013	1.15	0.250	-.3461729 1.330778
1.5 < 2	.3110203	.2983503	1.04	0.297	-.2737355 .8957762
2 < 2.5	.5522903	.2977427	1.85	0.064	-.0312747 1.135855
2.5 < 3	.706568	.2742707	2.58	0.010	.1690073 1.244129
3 < 3.5	.7836118	.2674484	2.93	0.003	.2594225 1.307801
3.5 < 4	.9235918	.2745488	3.36	0.001	.3854861 1.461698
4 < 4.5	.4578516	.2830636	1.62	0.106	-.0969428 1.012646
4.5 < 5	.7557408	.2767873	2.73	0.006	.2132477 1.298234
5 < 5.5	.4056986	.2660326	1.52	0.127	-.1157157 .9271129
5.5 < 6	.6968202	.2886221	2.41	0.016	.1311313 1.262509
6 < 6.5	1.021652	.318041	3.21	0.001	.3983035 1.645001
7 < 7.5	.3479793	.2906244	1.20	0.231	-.2216341 .9175927
7.5 < 8	.3969474	.2957657	1.34	0.180	-.1827426 .9766375
8 < 8.5	.8563053	.2855559	3.00	0.003	.296626 1.415985
8.5 < 9	.8936599	.3388502	2.64	0.008	.2295258 1.557794
9.5 < 10	.6660583	.2909823	2.29	0.022	.0957435 1.236373
10 < 11	.7880597	.326844	2.41	0.016	.1474572 1.428662
13 < 14	.7480036	.3285951	2.28	0.023	.1039691 1.392038
16 < 17	1.570876	.3153467	4.98	0.000	.952808 2.188944
_cons	1.263563	.2572284	4.91	0.000	.7594042 1.767721
/lnalpha	.430838	.0349712	12.32	0.000	.3622957 .4993803
siteid					
var(_cons)	.814985	.8342916			.1095932 6.060602

LR test versus nbinomial model: chibar2(01) = 0.00	Prob >= chibar2 = 1.0000
--	--------------------------

Table 4-20f: Protocol Workload (Protocols per FTE) on the Likelihood of Zero-Accruing Studies

Mixed-effects logistic regression	Number of obs	=	2,133
Group variable: siteid	Number of groups	=	7
	Obs per group:		
	min	=	101
	avg	=	304.7
	max	=	477
Integration method: mvaghermite	Integration pts.	=	7
Log likelihood = -1041.5903	Wald chi2(19)	=	26.14
	Prob > chi2	=	0.1265

nonaccr	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

prot_per_fte_cat						
1 < 1.5	.5517509	.6857201	0.80	0.421	-.7922358	1.895738
1.5 < 2	.118183	.5274036	0.22	0.823	-.915509	1.151875
2 < 2.5	.1002852	.540193	0.19	0.853	-.9584737	1.159044
2.5 < 3	.2054583	.4828128	0.43	0.670	-.7408374	1.151754
3 < 3.5	-.1874336	.472061	-0.40	0.691	-1.112656	.7377889
3.5 < 4	-.3562637	.4944209	-0.72	0.471	-1.325311	.6127834
4 < 4.5	-.3611947	.4927823	-0.73	0.464	-1.32703	.6046408
4.5 < 5	-.6850433	.5030105	-1.36	0.173	-1.670926	.3008391
5 < 5.5	-.3062167	.4677929	-0.65	0.513	-1.223074	.6106404
5.5 < 6	-.2122348	.5089708	-0.42	0.677	-1.209799	.7853297
6 < 6.5	-.8259901	.5872306	-1.41	0.160	-1.976941	.3249607
7 < 7.5	-.6188446	.5224803	-1.18	0.236	-1.642887	.405198
7.5 < 8	-.2589633	.5205591	-0.50	0.619	-1.27924	.7613138
8 < 8.5	-1.044586	.5489642	-1.90	0.057	-2.120536	.031364
8.5 < 9	-1.098818	.7349175	-1.50	0.135	-2.539229	.3415943
9.5 < 10	-.6731836	.528454	-1.27	0.203	-1.708934	.3625672
10 < 11	-.6632068	.6549164	-1.01	0.311	-1.946819	.6204057
13 < 14	-1.029664	.696594	-1.48	0.139	-2.394963	.335635
16 < 17	-.692283	.6376284	-1.09	0.278	-1.942012	.5574456
_cons	-.9447644	.5022484	-1.88	0.060	-1.929153	.0396244

siteid						
var(_cons)	.4012747	.2355191			.1270131	1.267755

LR test versus logistic model: chibar2(01) = 59.41				Prob >= chibar2 = 0.0000		

Table 4-20g: Protocol Workload (Protocols per FTE) on the Likelihood of Accruing at Least

Four Subjects

```
Mixed-effects logistic regression
Group variable:      siteid
Number of obs       =      2,133
Number of groups    =         7

Obs per group:
    min =         101
    avg =        304.7
    max =         477

Integration method: mvaghermite
Integration pts.    =         7

Log likelihood = -1410.3768
Wald chi2(19)      =         28.53
Prob > chi2        =         0.0737
```

actualaccrual_binary	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

prot_per_fte_cat						
1 < 1.5	.8602248	.6690053	1.29	0.199	-.4510014	2.171451
1.5 < 2	.1005062	.4693212	0.21	0.830	-.8193464	1.020359
2 < 2.5	.8433676	.4715677	1.79	0.074	-.080888	1.767623
2.5 < 3	.6017632	.4271686	1.41	0.159	-.2354718	1.438998
3 < 3.5	.7457778	.4142853	1.80	0.072	-.0662065	1.557762
3.5 < 4	.8525296	.4270287	2.00	0.046	.0155687	1.689491
4 < 4.5	.375649	.4393971	0.85	0.393	-.4855534	1.236851
4.5 < 5	1.055043	.4331179	2.44	0.015	.2061472	1.903938
5 < 5.5	.7410017	.417158	1.78	0.076	-.0766129	1.558616
5.5 < 6	.7075922	.4582246	1.54	0.123	-.1905115	1.605696
6 < 6.5	1.486202	.5095065	2.92	0.004	.4875875	2.484816
7 < 7.5	.972431	.4570222	2.13	0.033	.076684	1.868178
7.5 < 8	.6351451	.4669969	1.36	0.174	-.280152	1.550442
8 < 8.5	.6344372	.4575015	1.39	0.166	-.2622493	1.531124
8.5 < 9	.8805641	.5677754	1.55	0.121	-.2322553	1.993384
9.5 < 10	.7253916	.4834952	1.50	0.134	-.2222416	1.673025
10 < 11	1.126442	.5510871	2.04	0.041	.0463311	2.206553
13 < 14	1.287268	.5557346	2.32	0.021	.1980486	2.376488
16 < 17	1.048134	.5371535	1.95	0.051	-.004667	2.100936
_cons	-1.002871	.4649011	-2.16	0.031	-1.914061	-.0916818

siteid						
var(_cons)	.4112089	.2432563			.12898	1.311

LR test versus logistic model: chibar2(01) = 69.62			Prob >= chibar2 = 0.0000			

Table 4-20h: Protocol Workload (Protocols per FTE) on Activation Timeline

```
Mixed-effects ML regression
Group variable: siteid

Number of obs = 2,132
Number of groups = 7

Obs per group:
    min = 100
    avg = 304.6
    max = 477

Wald chi2(19) = 55.71
Prob > chi2 = 0.0000

Log likelihood = -6146.4823
```

activationtime_trans	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
-----+-----						
prot_per_fte_cat						
1 < 1.5	-.0478417	1.349067	-0.04	0.972	-2.691965	2.596282
1.5 < 2	1.653953	.9469871	1.75	0.081	-.2021072	3.510014
2 < 2.5	2.469629	.9505737	2.60	0.009	.6065384	4.332719
2.5 < 3	1.841462	.8629749	2.13	0.033	.1500626	3.532862
3 < 3.5	1.776327	.8346749	2.13	0.033	.1403941	3.41226
3.5 < 4	2.466825	.8606702	2.87	0.004	.7799428	4.153708
4 < 4.5	1.979444	.878785	2.25	0.024	.2570574	3.701831
4.5 < 5	1.313992	.8741238	1.50	0.133	-.3992593	3.027243
5 < 5.5	2.746416	.8364445	3.28	0.001	1.107015	4.385817
5.5 < 6	3.798996	.9271069	4.10	0.000	1.9819	5.616092
6 < 6.5	2.479842	1.030522	2.41	0.016	.4600557	4.499629
7 < 7.5	2.212435	.928403	2.38	0.017	.3927983	4.032071
7.5 < 8	2.592154	.947306	2.74	0.006	.7354686	4.44884
8 < 8.5	1.228306	.9255023	1.33	0.184	-.5856456	3.042257
8.5 < 9	.7074628	1.167257	0.61	0.544	-1.58032	2.995245
9.5 < 10	2.136391	.9525749	2.24	0.025	.2693786	4.003403
10 < 11	.0401613	1.126763	0.04	0.972	-2.168255	2.248577
13 < 14	.4770605	1.131758	0.42	0.673	-1.741144	2.695265
16 < 17	.5036093	1.097176	0.46	0.646	-1.646817	2.654035
_cons	10.50558	.9184733	11.44	0.000	8.705409	12.30576
-----+-----						

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
-----+-----				
siteid: Identity				
var(_cons)	1.510804	.8464948	.5038308	4.530347
-----+-----				
var(Residual)	18.49918	.5675219	17.41964	19.64563
-----+-----				

LR test versus linear model: chibar2(01) = 140.34 Prob >= chibar2 = 0.0000

Several variables modified the effect between protocols workload and clinical trial accrual. Factors that modified the protocols per staff member association include phase, IRB of record, inclusion of pediatric subjects, sponsor type, number of national sites, the total number of

months accruing nationally, the number of months accruing prior to local activation, and national enrollment goal. The inclusion of pediatric subjects and national enrollment goal only modified one category of the protocols per staff member variable (**Table 4-21**).

Table 4-21: Variable Association with Clinical Trial Accrual by Protocol Workload – All Sites

Variable	PROTOCOLS PER STAFF MEMBER	PROTOCOLS PER FULL-TIME EQUIVALENT (FTE)
Study phase	Effect Modifier	Effect Modifier
IRB of record	Effect Modifier	Effect Modifier
Sponsor type	Effect Modifier	Effect Modifier
Pediatric subjects	Effect Modifier	Not an Effect Modifier
Randomized	Not an Effect Modifier	Not an Effect Modifier
Placebo	Not an Effect Modifier	Not an Effect Modifier
Primary endpoint	Not an Effect Modifier	Not an Effect Modifier
National enrollment goal¥	Effect Modifier	Effect Modifier
# Sites nationally	Effect Modifier	Effect Modifier
Total # months accruing¥	Effect Modifier	Effect Modifier
# Months open prior to local open¥	Effect Modifier	Effect Modifier

Adjusted, reduced negative binomial regression models of the associated variable with protocol workload and outcome. Change in protocol workload beta coefficient $\geq 10\%$ were defined as “Effect Modifier.”

¥ Transformed variable used (log transformation for national enrollment and total number of months accruing; square root transformation for number of months of national accrual completed).

In addition, the following interaction terms were significant in reduced models of clinical trial accrual with protocols per staff member: national enrollment goal and number of national sites, national enrollment goal and total number of months accruing, and national enrollment goal and number of months accruing prior to local activation (**Table 4-22**).

Table 4-22: Effect Modification and Interaction Analyses by Protocol Workload – All Sites

VARIABLE INTERACTION PAIR WITH NATIONAL ENROLLMENT	# SITES	MONTHS NATIONAL ACCRUAL COMPLETED	TOTAL MONTHS PLANNED ACCRUAL	SPONSOR TYPE
CLINICAL TRIAL ACCRUAL				
PROTOCOLS PER STAFF MEMBER	Interaction	Effect Modifier	Interaction	Interaction
PROTOCOLS PER FTE	Interaction	Effect Modifier	Interaction	Interaction
ACTIVATION TIME				
PROTOCOLS PER STAFF MEMBER	Interaction	Interaction	Effect Modifier	Effect Modifier
PROTOCOLS PER FTE	Interaction	Interaction	Effect Modifier	Effect Modifier
ACTIVATION TIME				
PROTOCOLS PER STAFF MEMBER	Effect Modifier	Interaction	Effect Modifier	Interaction

Tests with an interaction term that had a corresponding p-value <0.05 defined as “Interaction.” If no interaction, models with a change in the beta coefficient of protocol workload $\geq 10\%$ were defined as “Effect Modifier,” otherwise listed as “Neither.”

The best fit full model included all the above-listed associated variables and interactions. These models were significant at a p-value of <0.001 and, after adjusting for all confounding factors, had a number of values in the protocol workload category that significantly and positively impacted clinical trial accrual (2.5 < 3.0, 3.0 < 3.5, 3.5 < 4.0, and 4.5 < 5.0 protocols per FTE; **Table 4-23**; 2.5 < 3.0, 3.5 < 4.0, and 5.5 < 6.0 protocols per staff member; **Table 4-24**).

Table 4-23: Full, Adjusted Model Assessing the Impact of Protocols per Staff Member on Clinical Trial Accrual

```

Mixed-effects nbinomial regression      Number of obs   =      2,133
Overdispersion:      mean
Group variable:      siteid             Number of groups =          7

Obs per group:
    min =      101
    avg =     304.7
    max =      477

Integration method: mvaghermite         Integration pts. =          7
Wald chi2(48) =     1027.72
Prob > chi2 =      0.0000
Log likelihood = -5733.6529
    
```

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

prot_per_staff_cat						
1 < 1.5	.3749936	.3602438	1.04	0.298	-.3310712 1.081058	
1.5 < 2	-.0448676	.2517506	-0.18	0.859	-.5382898 .4485546	
2 < 2.5	.4210857	.2392372	1.76	0.078	-.0478105 .8899819	
2.5 < 3	.4980758	.2297381	2.17	0.030	.0477974 .9483541	
3 < 3.5	.4364848	.225777	1.93	0.053	-.00603 .8789995	
3.5 < 4	.6182681	.2303985	2.68	0.007	.1666953 1.069841	
4 < 4.5	.3397038	.2336693	1.45	0.146	-.1182797 .7976873	
4.5 < 5	.3975086	.2322193	1.71	0.087	-.0576328 .85265	
5 < 5.5	.0888127	.2247583	0.40	0.693	-.3517055 .5293308	
5.5 < 6	.5463467	.2674359	2.04	0.041	.0221819 1.070511	
6 < 6.5	.3010109	.2509665	1.20	0.230	-.1908744 .7928961	
6.5 < 7	.136968	.2626588	0.52	0.602	-.3778339 .6517698	
7 < 7.5	-.035135	.2762989	-0.13	0.899	-.5766708 .5064008	
7.5 < 8	.0622508	.2690669	0.23	0.817	-.4651106 .5896122	
8 < 8.5	.4574952	.240825	1.90	0.057	-.0145131 .9295035	
8.5 < 9	.4696775	.283878	1.65	0.098	-.0867132 1.026068	
9.5 < 10	.2913607	.2440589	1.19	0.233	-.1869859 .7697074	
10 < 11	.4591922	.2738866	1.68	0.094	-.0776158 .9960001	
13 < 14	.5241908	.2766728	1.89	0.058	-.018078 1.06646	
16 < 17	.4057634	.2659002	1.53	0.127	-.1153915 .9269182	

phase						
Phase 0	-.1755798	.5891407	-0.30	0.766	-1.330274 .9791148	
Phase 1	-.0030458	.3080692	-0.01	0.992	-.6068504 .6007588	
Phase 1/2	.0472958	.312502	0.15	0.880	-.5651968 .6597884	
Phase 2	-.1021712	.3052554	-0.33	0.738	-.7004607 .4961184	

Phase 2/3	-.2286626	.3583019	-0.64	0.523	-.9309214	.4735962
Phase 3	-.290872	.3166556	-0.92	0.358	-.9115055	.3297615
Phase 4	.3547701	.4007185	0.89	0.376	-.4306237	1.140164
None	.1584294	.3472706	0.46	0.648	-.5222085	.8390672
irb						
Local	-.1334006	.0803279	-1.66	0.097	-.2908404	.0240392
sponsor_type						
Instititutional	-.2753947	.4683221	-0.59	0.557	-1.193289	.6424997
Industry	-.4638285	.4402541	-1.05	0.292	-1.326711	.3990537
National Group	-1.559212	.4887685	-3.19	0.001	-2.517181	-.6012435
peds						
Yes	-.1494688	.097456	-1.53	0.125	-.340479	.0415415
natenroll_trans	.8639371	.1482288	5.83	0.000	.5734139	1.15446
totalmo_trans	.5878265	.1539889	3.82	0.000	.2860137	.8896393
moaccrdone_trans	-.2958807	.0182316	-16.23	0.000	-.331614	-.2601475
natsitescat						
10-49 sites	-.6153816	.345921	-1.78	0.075	-1.293374	.062611
50-199 sites	-.3651807	.4094649	-0.89	0.372	-1.167717	.4373558
200+ sites	.7580302	.4834512	1.57	0.117	-.1895168	1.705577
Unknown	.6141845	.4811083	1.28	0.202	-.3287705	1.55714
natsitescat#c.natenroll_trans						
10-49 sites	-.0016987	.0825223	-0.02	0.984	-.1634394	.160042
50-199 sites	-.1415498	.0881016	-1.61	0.108	-.3142259	.0311262
200+ sites	-.350377	.0947372	-3.70	0.000	-.5360586	-.1646954
Unknown	-.2236427	.1095551	-2.04	0.041	-.4383668	-.0089186
c.natenroll_trans#c.totalmo_trans	-.0758445	.0321785	-2.36	0.018	-.1389131	-.0127759
sponsor_type#c.natenroll_trans						
Instititutional	.0895258	.1189922	0.75	0.452	-.1436946	.3227462
Industry	-.0457942	.106231	-0.43	0.666	-.2540031	.1624148
National Group	.1679912	.1128392	1.49	0.137	-.0531695	.3891519
_cons	-1.282324	.7366269	-1.74	0.082	-2.726086	.1614384

/lnalpha	-.1030164	.0411253	-2.50	0.012	-.1836206	-.0224122

siteid						
var(_cons)	.6160427	.6304525			.0828894	4.578495

7 < 7.5		.1176859	.2443639	0.48	0.630	-.3612585	.5966304
7.5 < 8		.1046975	.2490945	0.42	0.674	-.3835187	.5929137
8 < 8.5		.4469485	.2406321	1.86	0.063	-.0246817	.9185787
8.5 < 9		.461122	.2836673	1.63	0.104	-.0948556	1.0171
9.5 < 10		.2609232	.2438893	1.07	0.285	-.2170911	.7389375
10 < 11		.415787	.2731933	1.52	0.128	-.119662	.951236
13 < 14		.5136357	.2764076	1.86	0.063	-.0281132	1.055385
16 < 17		.3831712	.2659475	1.44	0.150	-.1380762	.9044187
phase							
Phase 0		-.1951483	.5885073	-0.33	0.740	-1.348601	.9583048
Phase 1		.0444669	.3081769	0.14	0.885	-.5595488	.6484826
Phase 1/2		.0786559	.3129439	0.25	0.802	-.5347027	.6920146
Phase 2		-.0530989	.3056268	-0.17	0.862	-.6521164	.5459187
Phase 2/3		-.1157736	.3573302	-0.32	0.746	-.8161279	.5845808
Phase 3		-.2466867	.3167084	-0.78	0.436	-.8674237	.3740503
Phase 4		.3604557	.4016951	0.90	0.370	-.4268522	1.147764
None		.21992	.3474493	0.63	0.527	-.4610681	.900908
irb							
Local		-.1305233	.0812744	-1.61	0.108	-.2898182	.0287716
sponsor_type							
Institutional		-.1187295	.4643266	-0.26	0.798	-1.028793	.7913339
Industry		-.3848782	.4355069	-0.88	0.377	-1.238456	.4686997
National Group		-1.5665	.4825849	-3.25	0.001	-2.512349	-.6206509
natenroll_trans		.8701608	.1476284	5.89	0.000	.5808144	1.159507
totalmo_trans		.5663316	.1532155	3.70	0.000	.2660348	.8666284
moaccrdone_trans		-.2950381	.0182205	-16.19	0.000	-.3307497	-.2593265
natsitescat							
10-49 sites		-.630429	.3451483	-1.83	0.068	-1.306907	.0460493
50-199 sites		-.3549685	.4078946	-0.87	0.384	-1.154427	.4444903
200+ sites		.9130522	.4802176	1.90	0.057	-.028157	1.854261
Unknown		.6120444	.4806778	1.27	0.203	-.3300668	1.554156
natsitescat#c.natenroll_trans							
10-49 sites		.0033794	.0824142	0.04	0.967	-.1581495	.1649082
50-199 sites		-.1445121	.0877106	-1.65	0.099	-.3164218	.0273975
200+ sites		-.3764295	.0941294	-4.00	0.000	-.5609197	-.1919392
Unknown		-.2311279	.1091122	-2.12	0.034	-.444984	-.0172719
c.natenroll_trans#c.totalmo_trans		-.0717852	.0319555	-2.25	0.025	-.1344168	-.0091537

sponsor_type#c.natenroll_trans							
Instititutional		.0597636	.1181087	0.51	0.613	-.1717252	.2912523
Industry		-.0631794	.1048385	-0.60	0.547	-.268659	.1423002
National Group		.1643959	.1112176	1.48	0.139	-.0535866	.3823784
	_cons	-1.354197	.7353317	-1.84	0.066	-2.795421	.0870269
/lnalpha		-.1021674	.0410546	-2.49	0.013	-.182633	-.0217018
siteid							
	var(_cons)	.5734352	.587823			.0768999	4.276053

LR test versus nbinomial model: $\text{chibar2}(01) = 0.00$ Prob >= $\text{chibar2} = 1.0000$

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	2,133	.	-5732.411	49	11562.82	11840.42

All examined independent variables with the exception of national enrollment goal independently acted as an effect modifier on the relationship between protocol workload and activation time (**Table 4-25**).

Table 4-25: Variable Association with Activation Time* by Protocol Workload – All Sites

Variable	PROTOCOLS PER STAFF MEMBER	PROTOCOLS PER FULL-TIME EQUIVALENT (FTE)
Study phase	Effect Modifier	Effect Modifier
IRB of record	Effect Modifier	Effect Modifier
Sponsor type	Effect Modifier	Effect Modifier
Pediatric subjects	Effect Modifier	Effect Modifier
Randomized	Effect Modifier	Effect Modifier
Placebo	Effect Modifier	Effect Modifier
Primary endpoint	Effect Modifier	Effect Modifier
National enrollment goal¥	Not an Effect Modifier	Not an Effect Modifier
# Sites nationally	Effect Modifier	Effect Modifier
Total # months accruing¥	Effect Modifier	Effect Modifier
# Months open prior to local open¥	Effect Modifier	Effect Modifier

Adjusted, reduced linear regression models of the associated variable with protocol workload and outcome. Change in protocol workload beta coefficient $\geq 10\%$ were defined as “Effect Modifier.”

**The dependent variable, activation time, is square root transformed to fit the observations to a normal distribution.*

¥ *Transformed variable used (log transformation for national enrollment and total number of months accruing; square root transformation for number of months of national accrual completed).*

However, in testing for interactions, national enrollment goal acted as an effect modifier with the terms total months of national accrual and sponsor type, therefore, it was also included in the final model. In addition, the interaction terms of national enrollment/number of sites as well as national enrollment goal/months of national enrollment already completed interacted with the association between protocol workload and activation time, therefore they were considered for inclusion in the full model (**Table 4-22**; page 195). In assessing best fit, only the interaction term of national enrollment goal and months of national enrollment already completed was included. The full, adjusted models assessing the impact between protocol workload and activation time were significant ($p < 0.001$) and are presented in **Tables 4-26** (protocols per staff member) and **4-27** (protocols per FTE). All workloads between 1.0 and 10 protocols per staff member or FTE (except for $8.0 < 8.5$) significantly increased activation times compared to the reference group of < 1 protocol per staff member.

Table 4-26: Full, Adjusted Model Assessing the Impact of Protocols per Staff Member on Activation Time

```

Mixed-effects ML regression      Number of obs   =    2,132
Group variable: siteid          Number of groups =         7

                                Obs per group:
                                min =    100
                                avg =   304.6
                                max =    477

                                Wald chi2(45)   =    809.56
                                Prob > chi2     =    0.0000

Log likelihood = -5831.1114
  
```

activationtime_trans	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

prot_per_staff_cat						
1 < 1.5	.074099	1.171443	0.06	0.950	-2.221887	2.370085
1.5 < 2	2.259864	.8231563	2.75	0.006	.6465073	3.873221
2 < 2.5	2.263456	.7814465	2.90	0.004	.7318487	3.795063
2.5 < 3	2.327481	.7519629	3.10	0.002	.8536609	3.801301
3 < 3.5	2.066682	.7366323	2.81	0.005	.6229098	3.510455
3.5 < 4	2.734754	.7497542	3.65	0.000	1.265263	4.204245
4 < 4.5	2.676973	.7547791	3.55	0.000	1.197633	4.156313
4.5 < 5	1.716394	.7660616	2.24	0.025	.2149411	3.217847
5 < 5.5	3.301516	.7308098	4.52	0.000	1.869155	4.733876
5.5 < 6	3.831552	.901381	4.25	0.000	2.064878	5.598226
6 < 6.5	3.111125	.8322619	3.74	0.000	1.479921	4.742328
6.5 < 7	2.045143	.8902464	2.30	0.022	.3002924	3.789994
7 < 7.5	3.073683	.9291807	3.31	0.001	1.252523	4.894844
7.5 < 8	3.462287	.9135637	3.79	0.000	1.671736	5.252839
8 < 8.5	1.454887	.8088237	1.80	0.072	-.1303781	3.040152
8.5 < 9	2.104052	1.020984	2.06	0.039	.1029588	4.105144
9.5 < 10	2.291233	.8307916	2.76	0.006	.6629112	3.919554
10 < 11	1.02532	.9847172	1.04	0.298	-.9046901	2.95533
13 < 14	.6650056	.988985	0.67	0.501	-1.273369	2.603381
16 < 17	1.216068	.9619516	1.26	0.206	-.6693221	3.101459

phase						
Phase 0	-7.228244	2.178025	-3.32	0.001	-11.4971	-2.959393
Phase 1	-4.728154	1.155398	-4.09	0.000	-6.992693	-2.463616
Phase 1/2	-4.411941	1.140043	-3.87	0.000	-6.646384	-2.177498
Phase 2	-4.75005	1.116291	-4.26	0.000	-6.937941	-2.56216
Phase 2/3	-4.950268	1.305991	-3.79	0.000	-7.509963	-2.390574
Phase 3	-4.426979	1.147261	-3.86	0.000	-6.67557	-2.178388

Phase 4	-6.365633	1.464688	-4.35	0.000	-9.23637	-3.494897
None	-5.352253	1.267143	-4.22	0.000	-7.835807	-2.868699
irb						
Local	1.360689	.3142131	4.33	0.000	.744843	1.976536
sponsor_type						
Institutitonal	1.052961	.4741984	2.22	0.026	.123549	1.982373
Industry	-.5378225	.4246654	-1.27	0.205	-1.370151	.2945063
National Group	-4.531673	.4589508	-9.87	0.000	-5.4312	-3.632146
peds						
Yes	-1.054073	.334265	-3.15	0.002	-1.70922	-.3989255
random						
Yes	.6672316	.2253791	2.96	0.003	.2254967	1.108967
placebo						
Yes	.2850586	.2680918	1.06	0.288	-.2403918	.8105089
priend						
Efficacy	.4790731	.353883	1.35	0.176	-.2145247	1.172671
Other	.2577464	.7771532	0.33	0.740	-1.265446	1.780939
natenroll_trans	-.5850067	.13468	-4.34	0.000	-.8489746	-.3210389
natsitescat						
10-49 sites	-.1091856	.2632005	-0.41	0.678	-.625049	.4066779
50-199 sites	-.045842	.3297676	-0.14	0.889	-.6921745	.6004906
200+ sites	.4419069	.4294863	1.03	0.304	-.3998709	1.283685
Unknown	-.4522076	.5198107	-0.87	0.384	-1.471018	.5666027
totalmo_trans	.3048293	.1649803	1.85	0.065	-.0185261	.6281847
moaccrdone_trans	-.0627828	.1904301	-0.33	0.742	-.436019	.3104534
natenroll_trans	0	(omitted)				
moaccrdone_trans	0	(omitted)				
c.natenroll_trans#c.moaccrdone_trans	.1725755	.0393178	4.39	0.000	.0955141	.2496369
_cons	13.90336	1.626645	8.55	0.000	10.71519	17.09152

```

-----
Random-effects Parameters | Estimate Std. Err. [95% Conf. Interval]
-----+-----
siteid: Identity         |
var(_cons) | 1.743665 .9776831 .5810178 5.232831
-----+-----
var(Residual) | 13.7425 .4216213 12.9405 14.59422
-----
LR test versus linear model: chibar2(01) = 149.78 Prob >= chibar2 = 0.0000

```

Akaike's information criterion and Bayesian information criterion

```

-----
Model | Obs ll(null) ll(model) df AIC BIC
-----+-----
. | 2,132 . -5831.111 48 11758.22 12030.13
-----

```

Table 4-27: Full, Adjusted Model Assessing Protocols per Full-Time Equivalent (FTE) on Activation Time

```

Mixed-effects ML regression      Number of obs   =   2,132
Group variable: siteid          Number of groups =     7
                                Obs per group:
                                min =   100
                                avg =  304.6
                                max =   477
                                Wald chi2(44)    =  805.98
                                Prob > chi2      =  0.0000
Log likelihood = -5832.2379

```

```

-----
activationtime_trans | Coef. Std. Err. z P>|z| [95% Conf. Interval]
-----+-----
prot_per_fte_cat |
1 < 1.5 | .1838545 1.172088 0.16 0.875 -2.113395 2.481104
1.5 < 2 | 2.249823 .8238376 2.73 0.006 .6351312 3.864515
2 < 2.5 | 2.655087 .8252798 3.22 0.001 1.037569 4.272606
2.5 < 3 | 2.131481 .7546577 2.82 0.005 .6523792 3.610583
3 < 3.5 | 2.230365 .7304142 3.05 0.002 .7987795 3.661951
3.5 < 4 | 2.66437 .7565367 3.52 0.000 1.181586 4.147155
4 < 4.5 | 2.438608 .7683676 3.17 0.002 .9326351 3.944581
4.5 < 5 | 1.839067 .764343 2.41 0.016 .340982 3.337152
5 < 5.5 | 3.130141 .7289529 4.29 0.000 1.70142 4.558862
5.5 < 6 | 4.072972 .8089033 5.04 0.000 2.48755 5.658393
6 < 6.5 | 3.532846 .9118759 3.87 0.000 1.745602 5.32009

```

7 < 7.5	2.803882	.8115924	3.45	0.001	1.213191	4.394574
7.5 < 8	3.221037	.829736	3.88	0.000	1.594785	4.84729
8 < 8.5	1.435891	.8081517	1.78	0.076	-.1480574	3.019839
8.5 < 9	2.159278	1.023187	2.11	0.035	.1538677	4.164688
9.5 < 10	2.317506	.8328928	2.78	0.005	.6850662	3.949946
10 < 11	1.082936	.9869896	1.10	0.273	-.8515275	3.017401
13 < 14	.736136	.9914469	0.74	0.458	-1.207064	2.679336
16 < 17	1.292621	.9645445	1.34	0.180	-.5978512	3.183094
phase						
Phase 0	-7.200975	2.174547	-3.31	0.001	-11.46301	-2.938941
Phase 1	-4.594401	1.155697	-3.98	0.000	-6.859526	-2.329277
Phase 1/2	-4.342461	1.14071	-3.81	0.000	-6.578212	-2.106709
Phase 2	-4.703476	1.116863	-4.21	0.000	-6.892487	-2.514465
Phase 2/3	-4.877179	1.306752	-3.73	0.000	-7.438366	-2.315992
Phase 3	-4.417251	1.146875	-3.85	0.000	-6.665085	-2.169416
Phase 4	-6.328469	1.466223	-4.32	0.000	-9.202213	-3.454725
None	-5.287539	1.268228	-4.17	0.000	-7.773221	-2.801858
irb						
Local	1.356559	.3153359	4.30	0.000	.738512	1.974606
sponsor_type						
Instituitonal	1.035686	.4743622	2.18	0.029	.1059532	1.965419
Industry	-.586908	.4248864	-1.38	0.167	-1.41967	.245854
National Group	-4.588234	.4592758	-9.99	0.000	-5.488398	-3.68807
peds						
Yes	-1.0409	.3347875	-3.11	0.002	-1.697072	-.3847291
random						
Yes	.6922548	.2250987	3.08	0.002	.2510694	1.13344
placebo						
Yes	.262799	.2682366	0.98	0.327	-.2629351	.7885331
priend						
Efficacy	.5158455	.3539479	1.46	0.145	-.1778796	1.209571
Other	.197234	.7772618	0.25	0.800	-1.326171	1.720639
natenroll_trans	-.5843889	.1347964	-4.34	0.000	-.848585	-.3201928
natsitescat						
10-49 sites	-.050721	.2633897	-0.19	0.847	-.5669553	.4655134
50-199 sites	.0246416	.3302296	0.07	0.941	-.6225964	.6718797

200+ sites		.5293993	.4303487	1.23	0.219	-.3140686	1.372867
Unknown		-.4962311	.5196071	-0.96	0.340	-1.514642	.5221802
totalmo_trans		.2744331	.1653205	1.66	0.097	-.0495892	.5984554
moaccrdone_trans		-.0409415	.1903518	-0.22	0.830	-.4140241	.3321411
natenroll_trans		0	(omitted)				
moaccrdone_trans		0	(omitted)				
c.natenroll_trans#c.moaccrdone_trans		.1696457	.0393276	4.31	0.000	.092565	.2467265
_cons		13.84064	1.61568	8.57	0.000	10.67396	17.00731

Random-effects Parameters		Estimate	Std. Err.	[95% Conf. Interval]	
siteid: Identity					
var(_cons)		1.452808	.8231139	.4785694	4.410333
var(Residual)		13.76506	.4223188	12.96172	14.61818

LR test versus linear model: $\chi^2(01) = 130.55$ Prob $\geq \chi^2 = 0.0000$

Akaike's information criterion and Bayesian information criterion

Model		Obs	ll(null)	ll(model)	df	AIC	BIC
.		2,132	.	-5832.238	47	11758.48	12024.72

Protocols per staff member, and not protocols per FTE, was found to be significantly associated with the likelihood to accrue at least four subjects (**Tables 4-20c** and **4-20g**; pages 188 and 192, respectively). Reduced models showed phase, sponsor type, inclusion of pediatric subjects, randomized design, primary endpoint, national enrollment goal, the number of national sites, and the number of months accruing nationally prior to local enrollment to be effect modifiers (**Table 4-28**). Total number of months accruing nationally, in combination with national enrollment goal, was also an effect modifier. National enrollment goal and the number of months of national enrollment already completed as well as national enrollment goal and sponsor type were interactors. However, in the full model, the best fit was without any interaction terms and keeping all terms as effect modifiers (**Table 4-29**).

Table 4-28: Variable Association with Likelihood to Accrue at Least Four Subjects by Protocol Workload – All Sites

Variable	PROTOCOLS PER STAFF MEMBER
Study phase	Effect Modifier
IRB of record	Not an Effect Modifier
Sponsor type	Effect Modifier
Pediatric subjects	Effect Modifier
Randomized	Effect Modifier
Placebo	Not an Effect Modifier
Primary endpoint	Effect Modifier
National enrollment goal¥	Effect Modifier

# Sites nationally	Effect Modifier
Total # months accruing¥	Not an Effect Modifier
# Months open prior to local open¥	Effect Modifier

Adjusted, reduced logistic regression models of the associated variable with protocol workload and outcome. Change in protocol workload beta coefficient $\geq 10\%$ were defined as “Effect Modifier.”

¥ Transformed variable used (log transformation for national enrollment and total number of months accruing; square root transformation for number of months of national accrual completed).

Table 4-29: Full, Adjusted Model Assessing Protocols per Staff Member on Likelihood of Accruing 4+ Subjects

```

Mixed-effects logistic regression      Number of obs   =      2,133
Group variable:      siteid           Number of groups =          7

Obs per group:
    min =          101
    avg =         304.7
    max =          477

Integration method: mvaghermite       Integration pts. =          7

Log likelihood = -1199.952             Wald chi2(42)   =      322.56
                                         Prob > chi2     =      0.0000
    
```

actualaccrual_binary	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

prot_per_staff_cat						
1 < 1.5	.8536866	.7218428	1.18	0.237	-.5610994 2.268473	
1.5 < 2	-.1256287	.5122513	-0.25	0.806	-1.129623 .8783654	
2 < 2.5	.8949191	.4911504	1.82	0.068	-.0677181 1.857556	
2.5 < 3	.6551404	.4710992	1.39	0.164	-.2681972 1.578478	
3 < 3.5	.746784	.4630863	1.61	0.107	-.1608485 1.654416	
3.5 < 4	.9283229	.4722371	1.97	0.049	.0027553 1.853891	
4 < 4.5	.3066582	.4749816	0.65	0.519	-.6242887 1.237605	
4.5 < 5	1.323449	.481545	2.75	0.006	.3796383 2.26726	
5 < 5.5	.605276	.4621508	1.31	0.190	-.3005229 1.511075	
5.5 < 6	1.025482	.5558431	1.84	0.065	-.0639507 2.114914	
6 < 6.5	1.36352	.526902	2.59	0.010	.3308111 2.396229	
6.5 < 7	1.07768	.5505732	1.96	0.050	-.0014237 2.156784	
7 < 7.5	.3638287	.5777693	0.63	0.529	-.7685784 1.496236	
7.5 < 8	.3624582	.5661909	0.64	0.522	-.7472555 1.472172	
8 < 8.5	.5142134	.5045354	1.02	0.308	-.4746577 1.503085	
8.5 < 9	.5076349	.6370287	0.80	0.426	-.7409185 1.756188	
9.5 < 10	.4262496	.5282779	0.81	0.420	-.609156 1.461655	
10 < 11	1.088412	.6128701	1.78	0.076	-.1127911 2.289616	
13 < 14	1.191477	.6243759	1.91	0.056	-.032277 2.415232	
16 < 17	.4673214	.5983656	0.78	0.435	-.7054536 1.640096	

phase						
Phase 0	.8435938	1.393153	0.61	0.545	-1.886935 3.574123	
Phase 1	1.333679	.7312098	1.82	0.068	-.099466 2.766824	
Phase 1/2	.9588834	.722789	1.33	0.185	-.4577571 2.375524	
Phase 2	.6691671	.7059354	0.95	0.343	-.7144409 2.052775	

Phase 2/3		-.2636982	.8262366	-0.32	0.750	-1.883092	1.355696
Phase 3		-.0881746	.7261546	-0.12	0.903	-1.511412	1.335062
Phase 4		1.36624	.9132331	1.50	0.135	-.423664	3.156144
None		.8494765	.81345	1.04	0.296	-.7448561	2.443809
sponsor_type							
Instituitonal		.0568541	.3169908	0.18	0.858	-.5644365	.6781446
Industry		-1.091172	.2783919	-3.92	0.000	-1.63681	-.5455341
National Group		-1.325378	.2967567	-4.47	0.000	-1.90701	-.7437454
peds							
Yes		-.5495624	.2090329	-2.63	0.009	-.9592594	-.1398654
random							
Yes		-.1826387	.1364066	-1.34	0.181	-.4499908	.0847134
friend							
Efficacy		.3405809	.2200861	1.55	0.122	-.0907799	.7719418
Other		.3797752	.4670029	0.81	0.416	-.5355337	1.295084
natenroll_trans		.7859568	.0673492	11.67	0.000	.6539547	.9179589
natsitescat							
10-49 sites		-.6591152	.1622842	-4.06	0.000	-.9771864	-.3410441
50-199 sites		-1.066812	.2040201	-5.23	0.000	-1.466684	-.6669397
200+ sites		-1.097687	.2680704	-4.09	0.000	-1.623095	-.5722784
Unknown		.0209758	.3277652	0.06	0.949	-.6214322	.6633839
moaccrdone_trans		-.4305276	.0397557	-10.83	0.000	-.5084473	-.3526079
totalmo_trans		.2038198	.1008459	2.02	0.043	.0061655	.4014742
_cons		-3.359368	.9725106	-3.45	0.001	-5.265454	-1.453282

siteid							
var(_cons)		.3941148	.2384327			.1204094	1.289986

LR test versus logistic model: $\chi^2(01) = 48.95$ Prob $\geq \chi^2 = 0.0000$

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	2,133	.	-1199.952	44	2487.904	2737.177

4.5.5 Optimal Protocol Workload for Clinical Trial Accrual

The overall distribution of clinical trial accrual associated with protocol workload is presented in **Figure 4-5** (protocols per staff member; page 175) and **Figure 4-6** (protocols per FTE; page 176). To get a better view of the distribution of the medians and quartiles, accrual was also plotted limiting the studies to those that accrued under 100 subjects (plots **Figure 4-5b** and **4-6b**). Optimal workload for maximizing clinical trial accrual at individual sites varied from < 1 protocol per staff member/FTE to $8.0 < 8.5$ protocol per staff member/FTE. At the site level, the most common protocol workloads that maximized median accrual were $5.0 < 5.5$ and $6.0 < 6.5$ protocols per staff member and $2.5 < 3.0$, $4.5 < 5.0$, and $5.0 < 5.5$ protocols per FTE (**Tables 4-30** and **4-31**).

Table 4-30: Best Workload Performance by Study Characteristic Analysis – Protocols per Staff Member

SITE	LOWEST ZERO-ACCRUING (%)	HIGHEST 4+ ACCRUING (%)	MEDIAN TRIAL ACCRUAL	LOWEST MEDIAN ACTIVATION TIME
ALL	8.5 < 9.0 (11.8%)	13 < 14 (60.0%)* 2 < 2.5 (57.9%)	16 < 17 (4.5)* 2.0 < 2.5; 2.5 < 3.0; 3.0 < 3.5; 3.5 < 4.0; 4.5 < 5.0; 6.0 < 6.5 (4)	1.0 < 1.5 (96)
448155	2.5 < 3.0 (5.4%)	3.5 < 4.0 (58.9%)	2.5 < 3.0 (4)	3.0 < 3.5 (121)
494048	6.0 < 6.5 (8.3%)	3.0 < 3.5 (68.4%)	6.0 < 6.5 (9)	< 1 (47)** 2.5 < 3.0 (57)
512786	5.0 < 5.5 (4.0%)	2.0 < 2.5 (73.9%)	5.0 < 5.5 (8)	1.5 < 2.0 (118)
560623	4.0 < 4.5 (20.8%)	4.5 < 5.0 (33.3%)	4.0 < 4.5 (1.5)	2.0 < 2.5 (72.5)
602591	6.0 < 6.5 (12.5%)	6.0 < 6.5 (56.3%)	2.5 < 3.0; 6.0 < 6.5 (4)	< 1 (41)
696337	< 1 (0.0%)	< 1 (57.1%)	< 1; 8.0 < 8.5 (4)	8.0 < 8.5 (20)
714415	8.5 < 9.0 (11.8%)	5.0 < 5.5 (70.0%)	5.0 < 5.5 (5)	< 1 (59.5)** 10 < 11 (97)

* Only one institution reported this level of staffing. Therefore, second best performing metric also presented.

** Only two studies in this category; second best metric also presented.

Table 4-31: Best Workload Performance by Study Characteristic Analysis – Protocols per FTE

SITE	LOWEST ZERO-ACCRUING (%)	HIGHEST 4+ ACCRUING (%)	MEDIAN TRIAL ACCRUAL	LOWEST MEDIAN ACTIVATION TIME
ALL	8.5 < 9.0 (11.8%)	13 < 14 (60.0%)* 2 < 2.5 (57.9%)	16 < 17 (4.5)* 2.0 < 2.5; 2.5 < 3.0; 3.0 < 3.5; 3.5 < 4.0; 4.5 < 5.0; 6.0 < 6.5 (4)	4.5 < 5.0 (94.5)
448155	3.5 < 4.0 (10.5%)	3.5 < 4.0 (58.9%)	2.5 < 3.0; 3.0 < 3.5; 3.5 < 4.0; 4.5 < 5.0 (4)	1.5 < 2.0 (122)
494048	6.0 < 6.5 (8.3%)	3.0 < 3.5 (68.4%)	6.0 < 6.5 (9)	< 1 (47)** 2.5 < 3.0 (57)
512786	5.0 < 5.5 (4.0%)	2.0 < 2.5 (73.9%)	5.0 < 5.5 (8)	1.5 < 2.0 (118)
560623	4.0 < 4.5 (20.8%)	4.5 < 5.0 (33.3%)	4.0 < 4.5 (1.5)	2.0 < 2.5 (72.5)
602591	<1 (0.0%)	2.5 < 3.0 (53.9%)	2.5 < 3.0; 7.0 < 7.5 (4)	< 1 (41)
696337	<1 (0.0%)	<1 (57.1%)	<1; 8.0 < 8.5 (4)	8.0 < 8.5 (20)
714415	8.5 < 9.0 (11.8%)	5.0 < 5.5 (70.0%)	5.0 < 5.5 (5)	< 1 (59.5)** 10 < 11 (97)

* Only one institution reported this level of staffing. Therefore, second best performing metric also presented.

** Only two studies in this category; second best metric also presented.

Combining all site data, the highest median accrual per protocol was at a staffing level of $16 < 17$ protocols per staff member, at 17 subjects per protocol (median 4.5); however this data was submitted by a single institution. For protocol workloads under 10 protocols per staff member/FTE, several staffing levels had a median subject accrual of four subjects ($2.0 < 2.5$, $2.5 < 3.0$, $3.0 < 3.5$, $3.5 < 4.0$, $4.5 < 5.0$, $6.0 < 6.5$). The protocol workload that had the highest average accrual within this group was $3.5 < 4.0$ protocols per staff member (average 8.4 subjects per protocol; **Table 4-18**; page 181) and $6.0 < 6.5$ protocols per FTE (average 8.9 subjects per protocol; **Table 4-19**; page 183). Per the regression analysis, the number of protocols per staff member that corresponded with the largest gain to clinical trial accrual was $3.5 < 4.0$, with a gain of 1.86 accruals, holding all other variables constant (**Table 4-23**; page 197). Protocols per FTE also had the largest gain to clinical trial accrual at the workload level of $3.5 < 4.0$, at 1.88 accruals, holding all other variables constant (**Table 4-24**; page 199).

Compensating for the potential of not having the complete activation workload, sensitivity analyses were performed. With a 10% increase, optimal protocol workload increases to $3.9 < 4.4$ protocols per staff member (study characteristic analysis and regression analysis) and FTE (regression analysis), and $6.6 < 7.2$ protocols per FTE (study characteristic analysis). With a 25% increase in protocol workload, the number of protocols per staff member that maximizes study accrual increases to $4.4 < 5.0$ (study characteristic analysis and regression analysis) and FTE (regression analysis). The number of protocols per FTE increases to $7.5 < 8.1$ protocols (study characteristic analysis). The range of optimal protocol workload to maximize clinical trial

accrual is $3.5 < 5.0$ protocols per staff member (study characteristic and regression analyses) and $3.5 < 5.0$ (regression analysis) or $6.0 < 8.1$ protocols per FTE (study characteristic analysis).

4.5.6 Optimal Protocol Workload for Zero-Accruing Protocols

Within this Specific Aim population, 42 (20.7%) of studies did not accrue any subjects (**Tables 4-18** and **4-19**; pages 181 and 183, respectively). Optimal workload minimizing the percentage of protocols that do not accrue any subjects at individual sites varied from < 1 protocol per staff member/FTE to $8.5 < 9.0$ protocols per staff member/FTE. At the site level, the most common protocol workloads that minimized zero-accruing protocols were $6.0 < 6.5$ protocols per staff member and < 1 protocol per FTE (**Tables 4-30** and **4-31**; pages 214 and 215, respectively).

Combining all site data, the lowest percentage of zero-accruing trials was 11.8% at the $8.5 < 9.0$ protocols per staff member (**Table 4-18**; page 181) and FTE (**Table 4-19**; page 183).

Compensating for the potential of not having the complete activation workload, sensitivity analyses were performed. With a 10% increase, protocol workload increases to $9.4 < 9.9$ protocols per staff member/FTE. With a 25% increase in protocol workload, the number of protocols per staff member/FTE that minimizes zero-accruing studies increases to $10.6 < 11.3$.

The range of optimal protocol workload that minimizes zero-accruing studies is $8.5 < 11.3$ protocols per staff member/FTE. Regression modeling showed no significant association between either protocol workload measure, protocols per staff member or protocols per FTE, and the likelihood that a study would accrue zero subjects.

4.5.7 Optimal Protocol Workload for Four-Plus Accruing Studies

Within this Specific Aim population, 1,012 (47.4%) of studies accrued at least four subjects (**Tables 4-18** and **4-19**; pages 181 and 183, respectively). Optimal workload maximizing the percentage of protocols that accrue at least four subjects at individual sites varied from < 1 to $6.0 < 6.5$ protocols per staff member and < 1 to $5.0 < 5.5$ protocols per FTE. At the site level, the most common protocol workloads that maximized protocols that accrue at least four subjects were $6.0 < 6.5$ protocols per staff member and < 1 protocol per FTE (**Tables 4-30** and **4-31**; pages 214 and 215, respectively).

Combining all site data, the highest percentage of four-plus accruing trials was 60.0% at the $13 < 14$ protocols per staff member (**Table 4-18**; page 181) and FTE (**Table 4-19**; page 183); however this data was submitted by a single institution. Examining workloads under 10 protocols per staff member/FTE, the protocol workload that maximized the percentage of studies that accrue at least four subjects was $2.0 < 2.5$ protocols per staff member/FTE (study characteristic analyses). There was no statistically significant relationship between the likelihood that a study would accrue at least four subjects and protocol workload per FTE; however, there was a statistically significant relationship with the number of protocols per staff member ($p=0.002$). The protocol workload that most increased the odds of accruing at least four subjects was $4.5 < 5.0$ subjects.

Compensating for the potential of not having the complete activation workload, sensitivity analyses were performed. With a 10% increase, protocol workload increases to $2.2 < 2.8$ protocols per staff member/FTE for the study characteristic analyses and $5.0 < 5.5$ protocols per

staff member by regression analysis. With a 25% increase in protocol workload, the range of optimal protocol workload to maximize four-plus accruing protocols is $2.0 < 3.1$ protocols per staff member/FTE for study characteristic analyses and $5.6 < 6.3$ protocols per staff member by regression analysis. The optimal workload to maximize the likelihood that a study will accrue at least four subjects is $2.0 < 3.1$ protocols per staff member/FTE (study characteristic analyses) and $4.5 < 6.3$ protocols per staff member by regression analysis.

4.5.8 Optimal Protocol Workload for Activation Timeline

The distribution of activation times associated with protocol workload are presented in **Figure 4-3** (page 143). Only considering categories that had more than two studies contributing to the metric, optimal workload for minimizing activation times at individual sites varied from < 1 protocol per staff member/FTE to $10 < 11$ protocol per staff member/FTE. At the site level, the most common protocol workloads that minimized activation time was < 1 and $1.5 < 2.0$ protocols per FTE (**Table 4-31**; page 215). There was no category of protocol workload per staff member that produced the lowest median activation time at more than one institution (**Table 4-30**; page 214).

Combining all site data, the lowest median activation time per protocol was at a staffing level of $1.0 < 1.5$ protocols per staff member and $4.5 < 5.0$ protocols per FTE (**Tables 4-18** and **4-19**; pages 181 and 183, respectively). Per regression analyses, the number of protocols per staff member that corresponded with the lowest clinical trial activation times was $1.0 < 1.5$, with an associated value of 1.1 added days to the square root of activation time compared to the majority

of workloads, which added a minimum of 7.4 days to the square root of activation time, holding all other variables constant (**Table 4-26**; page 204). Protocols per FTE also had the smallest associated value of added days at the workload level of $1.0 < 1.5$ protocols per FTE, with a gain of 1.2 days, holding all other variables constant (**Table 4-27**; page 206).

Compensating for the potential of not having the complete activation workload, sensitivity analyses were performed. With a 10% increase, protocol workload increases to $1.1 < 1.7$ protocols per staff member (study characteristic analyses and regression analyses) and $5.0 < 5.5$ protocols per FTE (study characteristic analyses) and $1.1 < 1.7$ protocols per FTE (regression analyses). With a 25% increase in protocol workload, the number of protocols per staff member that maximizes study accrual increases to $1.3 < 1.9$ (study characteristic analyses and regression analysis). The number of protocols per FTE increases to $5.6 < 6.3$ protocols (study characteristic analyses) and $1.3 < 1.9$ (regression analysis). Therefore, the range of optimal protocol workload to minimize median activation time is between 1.0 and 1.9 protocols per staff member (and per FTE, regression analysis) and 4.5 and 6.3 protocols per FTE (study characteristic analysis).

4.5.9 Site-Specific Analyses

Univariate analyses showed varying association between protocol workload and accrual outcomes (trial accrual, likelihood of zero-accruing studies, likelihood of accruing at least four subjects) and activation time. No single variable was consistently significant at all sites (**Table 4-32**). At site 448155, protocol workload (protocols per staff member or protocols per FTE) was

not associated with any outcome. Protocol workload was statistically significantly related to all outcomes for site 494048.

Table 4-32: Site-Specific Protocol Workload Univariate Analyses

SITE	ZERO-ACCRUING		4+ ACCRUING		ACCRUAL		ACTIVATION TIME	
	per Staff Member	per FTE	per Staff Member	per FTE	per Staff Member	per FTE	per Staff Member	per FTE
448155	0.09	0.24	0.54	0.48	0.06	0.06	0.70	0.54
494048	0.03	0.03	<0.001	<0.001	<0.001	<0.001	0.004	0.004
512786	0.24	0.24	0.007	0.007	0.06	0.06	0.07	0.07
560623	0.008	0.008	0.24	0.24	0.01	0.01	0.44	0.44
602591	0.03	0.03	0.07	0.10	0.002	0.001	<0.001	<0.001
696337	0.40	0.40	0.85	0.85	0.23	0.23	<0.001	<0.001
714145	0.79	0.79	0.34	0.34	<0.001	<0.001	0.06	0.06

P-values of univariate analyses between protocol workload (protocols per staff member or FTE) and efficiency outcome (likelihood to accrue zero subjects, likelihood to accrue at least four subjects, local accrual, activation time). Significant values ($p < 0.05$) are bolded.

4.5.9.1 Site-Specific Analysis: Protocol Workload

In site-specific models, protocol workload was significantly associated with clinical trial accrual at four sites (494048, 560623, 602591, and 714145). Protocol workload was associated with zero-accruing trials at three sites (494048, 560623, and 602591) and protocols accruing at least four subjects at two sites (494048 and 512786). Protocol workload was associated with

activation time at three sites (494048, 602591, and 696337). Variable association and interaction analyses are presented in **Tables 4-33** through **4-37**. Final, full models included variables with the strongest associations (shown by extreme z-values) with a total of no more degrees of freedom than 15% of the total number of observations. Final, full models are presented in **Appendix F-8** (clinical trial accrual), **Appendix F-9** (zero-accruing trials), **Appendix F-10** (trials accruing at least four subjects), and **Appendix F-11** (activation time).

Table 4-33: Independent Variable Association of Protocol Workload with Clinical Trial Accrual by Site

Variable	SITE 494048		SITE 560623		SITE 602591		SITE 714415	
	Per Staff Member	Per FTE	Per Staff Member	Per FTE	Per Staff Member	Per FTE	Per Staff Member	Per FTE
Study phase	Effect Modifier		Effect Modifier		Effect Modifier	Not an Effect Modifier	Effect Modifier	
IRB of record	*		Not an Effect Modifier		Effect Modifier	Effect Modifier	Effect Modifier	
Sponsor type	Effect Modifier		Effect Modifier		Effect Modifier	Effect Modifier	Effect Modifier	
Pediatric subjects	Effect Modifier		Effect Modifier		Effect Modifier	Not an Effect Modifier	Effect Modifier	
Randomized	Not an Effect Modifier		Effect Modifier		Not an Effect Modifier	Effect Modifier	Effect Modifier	
Placebo	Effect Modifier		Effect Modifier		Not an Effect Modifier	Not an Effect Modifier	Not an Effect Modifier	
Primary endpoint	Not an Effect Modifier		Effect Modifier		Not an Effect Modifier	Not an Effect Modifier	Effect Modifier	
National enrollment goal¥	Effect Modifier		Effect Modifier		Effect Modifier	Effect Modifier	Effect Modifier	
# Sites nationally	Effect Modifier		Effect Modifier		Effect Modifier	Effect Modifier	Effect Modifier	

Total # months accruing ¥	Effect Modifier	Effect Modifier	Effect Modifier	Effect Modifier	Not an Effect Modifier
# Months open prior to local open ¥	Effect Modifier	Effect Modifier	Effect Modifier	Effect Modifier	Effect Modifier
Treatment study	Effect Modifier	Effect Modifier	Not an Effect Modifier	Not an Effect Modifier	Effect Modifier
Precision medicine	Not an Effect Modifier	Effect Modifier	Effect Modifier	Effect Modifier	Not an Effect Modifier
Disease Team	Effect Modifier	Effect Modifier	Effect Modifier	Effect Modifier	Effect Modifier

Reduced negative binomial regression models of the associated variable with protocol workload and outcome. Change in protocol workload beta coefficient $\geq 10\%$ were defined as “Effect Modifier.”

** Collinear due to the institution only utilizing one IRB.*

*** Data not provided.*

¥ Transformed variable used (log transformation for national enrollment and total number of months accruing; square root transformation for number of months of national accrual completed).

Table 4-34: Independent Variable Association of Protocol Workload with Zero-Accruing Protocols by Site

Variable	SITE 494048		SITE 560623		SITE 602591	
	Per Staff Member	Per FTE	Per Staff Member	Per FTE	Per Staff Member	Per FTE
Study phase	Effect Modifier		Effect Modifier		Effect Modifier	Effect Modifier
IRB of record	*		Effect Modifier		*	*
Sponsor type	Effect Modifier		Effect Modifier		Effect Modifier	Effect Modifier
Pediatric subjects	Not an Effect Modifier		Not an Effect Modifier		Effect Modifier	Effect Modifier
Randomized	Effect Modifier		Effect Modifier		Not an Effect Modifier	Not an Effect Modifier
Placebo	Effect Modifier		Not an Effect Modifier		Not an Effect Modifier	Not an Effect Modifier
Primary endpoint	Effect Modifier		Effect Modifier		Not an Effect Modifier	Not an Effect Modifier
National enrollment goal¥	Effect Modifier		Effect Modifier		Effect Modifier	Effect Modifier
# Sites nationally	Effect Modifier		Effect Modifier		Effect Modifier	Effect Modifier

Total # months accruing ¥	Effect Modifier	Effect Modifier	Effect Modifier	Effect Modifier
# Months open prior to local open ¥	Effect Modifier	Effect Modifier	Effect Modifier	Effect Modifier
Treatment study	Not an Effect Modifier	Effect Modifier	Not an Effect Modifier	Not an Effect Modifier
Precision medicine	Effect Modifier	Effect Modifier	Not an Effect Modifier	Not an Effect Modifier
Disease Team	Effect Modifier	Effect Modifier	Effect Modifier	Effect Modifier

Reduced logistic regression models of the associated variable with protocol workload and outcome. Change in protocol workload beta coefficient $\geq 10\%$ were defined as “Effect Modifier.”

** Collinear due to the institution only utilizing one IRB.*

¥ Transformed variable used (log transformation for national enrollment and total number of months accruing; square root transformation for number of months of national accrual completed).

Table 4-35: Independent Variable Association of Protocol Workload with At Least Four Accruals by Site

Variable	SITE 494048		SITE 512786	
	Per Staff Member	Per FTE	Per Staff Member	Per FTE
Study phase	Effect Modifier		Effect Modifier	
IRB of record	*		Effect Modifier	
Sponsor type	Effect Modifier		Effect Modifier	
Pediatric subjects	Effect Modifier		Not an Effect Modifier	
Randomized	Effect Modifier		Effect Modifier	
Placebo	Effect Modifier		Effect Modifier	
Primary endpoint	Effect Modifier		Effect Modifier	
National enrollment goal $\sqrt{\text{¥}}$	Effect Modifier		Not an Effect Modifier	
# Sites nationally	Effect Modifier		Effect Modifier	
Total # months accruing $\sqrt{\text{¥}}$	Effect Modifier		Not an Effect Modifier	
# Months open prior to local open $\sqrt{\text{¥}}$	Effect Modifier		Effect Modifier	
Treatment study	Effect Modifier		**	
Precision medicine	Effect Modifier		**	
Disease Team	Effect Modifier		Effect Modifier	

Reduced logistic regression models of the associated variable with protocol workload and outcome. Change in protocol workload beta coefficient $\geq 10\%$ were defined as “Effect Modifier.”

** Collinear due to the institution only utilizing one IRB** Data not provided*

$\sqrt{\text{¥}}$ Transformed variable used (log transformation for national enrollment and total number of months accruing; square root transformation for number of months of national accrual completed).

Table 4-36: Independent Variable Association of Protocol Workload with Activation Time by Site

Variable	SITE 494048		SITE 602591		SITE 696337	
	Per Staff Member	Per FTE	Per Staff Member	Per FTE	Per Staff Member	Per FTE
Study phase	Effect Modifier		Effect Modifier	Effect Modifier	Effect Modifier	
IRB of record	*		Effect Modifier	Effect Modifier	*	
Sponsor type	Effect Modifier		Effect Modifier	Effect Modifier	Effect Modifier	
Pediatric subjects	Effect Modifier		Not an Effect Modifier	Not an Effect Modifier	Effect Modifier	
Randomized	Not an Effect Modifier		Not an Effect Modifier	Not an Effect Modifier	Effect Modifier	
Placebo	Effect Modifier		Not an Effect Modifier	Not an Effect Modifier	Effect Modifier	
Primary endpoint	Effect Modifier		Not an Effect Modifier	Not an Effect Modifier	Not an Effect Modifier	
National enrollment goal¥	Effect Modifier		Not an Effect Modifier	Effect Modifier	Effect Modifier	
# Sites nationally	Effect Modifier		Effect Modifier	Effect Modifier	Effect Modifier	

Total # months accruing¥	Effect Modifier	Not an Effect Modifier	Effect Modifier	Not an Effect Modifier
# Months open prior to local open¥	Effect Modifier	Effect Modifier	Effect Modifier	Not an Effect Modifier
Treatment study	Not an Effect Modifier			
Precision medicine	Effect Modifier	Effect Modifier	Effect Modifier	Not an Effect Modifier
Disease Team	Effect Modifier	Effect Modifier	Effect Modifier	Effect Modifier

Reduced linear regression models of the associated variable with protocol workload and and outcome. Change in protocol workload beta coefficient $\geq 10\%$ were defined as “Effect Modifier.”

** Collinear due to the institution only utilizing one IRB.*

*** Data not provided.*

¥ Transformed variable used (log transformation for national enrollment and total number of months accruing; square root transformation for number of months of national accrual completed).

Table 4-37: Effect Modification and Interaction Analyses for Protocol Workload by Site

INTERACTION PAIR WITH NATIONAL ENROLLMENT	# SITES		MONTHS NATIONAL ACCRUAL COMPLETED		TOTAL MONTHS ACCRUING		SPONSOR TYPE	
	Per Staff Member	Per FTE	Per Staff Member	Per FTE	Per Staff Member	Per FTE	Per Staff Member	Per FTE
CLINICAL TRIAL ACCRUAL								
SITE 494048	Effect Modifier		Effect Modifier		Effect Modifier		Effect Modifier	
SITE 560623	Effect Modifier		Effect Modifier		Effect Modifier		Effect Modifier	
SITE 602591	Effect Modifier	Effect Modifier	Effect Modifier	Effect Modifier	Interaction	Interaction	Effect Modifier	Effect Modifier
SITE 714415	Interaction		Interaction		Interaction		Interaction	
ZERO-ACCRUING TRIALS								
SITE 494048	Effect Modifier		Effect Modifier		Effect Modifier		Effect Modifier	
SITE 560623	Effect Modifier		Effect Modifier		Effect Modifier		Not computable	
SITE 602591	Effect Modifier	Effect Modifier	Effect Modifier	Effect Modifier	Interaction	Interaction	Effect Modifier	Effect Modifier
4+ ACCRUING TRIALS								
SITE 494048	Interaction		Effect Modifier		Effect Modifier		Interaction	

SITE 512786	Effect Modifier	Effect Modifier	Effect Modifier	Interaction				
ACTIVATION TIME								
SITE 494048	Effect Modifier	Interaction	Effect Modifier	Interaction				
SITE 602591	Interaction	Interaction	Interaction	Interaction	Interaction	Interaction	Effect Modifier	Effect Modifier
SITE 696337	Interaction	Interaction	Interaction	Interaction	Interaction	Interaction	Effect Modifier	Effect Modifier

Tests with an interaction term that had a corresponding p-value <0.05 defined as “Interaction.” If no interaction, models with a change in the beta coefficient of protocol workload $\geq 10\%$ were defined as “Effect Modifier,” otherwise listed as “Neither.”

Optimal protocol workload varied both by site and by outcome (**Table 4-38**). Activation time was minimized with a protocol workload of < 2.0 protocols per staff member/FTE at all three sites. At between 3.0 and 4.0 protocols per staff member/FTE, the likelihood that protocols would accrue at least four subjects was maximized (two sites). Above 4.0 protocols per staff member/FTE was optimal protocol workload for minimizing zero-accruing protocols (range 4.0 < 4.5 to 7.0 < 7.5) and maximizing clinical trial accrual (4.5 < 5.0 to 13 < 14).

Table 4-38: Optimal Protocol Workload by Regression Analysis

SITE	# ZERO-ACCRUING (%)		# 4+ ACCRUING (%)		ACCRUAL		ACTIVATION TIME	
	per Staff Member	per FTE	per Staff Member	per FTE	per Staff Member	per FTE	per Staff Member	per FTE
ALL	NS	NS	4.5 < 5.0	NS	3.5 < 4.0	3.5 < 4.0	1.0 < 1.5	1.0 < 1.5
448155	NS		NS		NS		NS	
494048	6.0 < 6.5		3.0 < 3.5		5.5 < 6.0		1.0 < 1.5	
512786	NS		3.5 < 4.0		NS		NS	
560623	4.0 < 4.5		NS		4.5 < 5.0		NS	
602591	6.0 < 6.5	7.0 < 7.5	NS		6.5 < 7.0	7.0 < 7.5	< 1	< 1
696337	NS		NS		NS		1.5 < 2.0	
714145	NS		NS		13 < 14		NS	

NS: not significant in the univariate regression analysis

4.5.10 Specific Aim Two Summary

- Hypothesis 2a Finding: **Reject the null hypothesis** ($p < 0.001$)
- Hypothesis 2b Finding: **Reject the null hypothesis** ($p < 0.001$)
- Hypothesis 2c Finding: **Fail to reject the null hypothesis** ($p = 0.13$)
- Hypothesis 2d Finding: **Reject the null hypothesis** ($p < 0.001$)
- Hypothesis 2e Finding: **Reject the null hypothesis** ($p < 0.001$)
- Hypothesis 2f Finding: **Fail to reject the null hypothesis** ($p = 0.13$).

Protocol workload did not act as an inverse curvilinear function, as described by the Multiple Team Membership (MTM) framework, but instead was better described as a categorical variable. A varying range of workload values were associated with better performance depending on the outcome of interest. The protocol workloads that best optimized the outcome of interest are summarized in **Tables 4-30** and **4-31** (study characteristic analyses; pages 214 and 215, respectively) and **Table 4-38** (regression analyses; page 232). To maximize clinical trial accrual, three of the four analyses showed an optimal workload of $3.5 < 5.0$ protocols. While there was no significant relationship between protocol workload (either by staff member or FTE) and the likelihood that a study would accrue no subjects, the protocol workload that had the lowest percentage of zero-accurring studies was $8.5 < 11.3$ protocols per staff member or FTE. To maximize the number of protocols that accrued at least four subjects, study characteristic analyses found the optimal workload to be between 2.0 and 3.1 protocols per staff member or FTE. Regression analysis showed no significant relationship between protocols per FTE, but that a protocol per staff member workload of between 4.5 and 6.3 protocols per staff member

maximized the likelihood that a study would accrue at least four subjects. Finally, protocol workload was examined to determine what workload minimized activation time. The optimal workload was found to be between 1.0 and 1.9 protocols per staff member (all analyses) or FTE (by regression analysis).

4.6 Specific Aim Three

4.6.1 Univariate Analyses

Each independent variable was assessed at the site level using univariate negative binomial regression models for their association with local accrual (**Table-4-39**). There was a significant association between each variable and local accrual at every institution. Only one variable was significantly associated with local accrual at all 16 participating institutions, the number of national sites (categorical). Variables significant at 15 of the 16 institutions include sponsor type, total months of national accrual, and the number of months of national accrual prior to local opening. Significant at 14 of 16 institutions was the variable of disease team. One variable was significant at only one institution. This variable was precision medicine, significant only at site 846594 ($p=0.02$).

Table 4-39: Independent Variable Association with Clinical Trial Accrual by Site

Variable	SITE 104647	SITE 173472	SITE 448155	SITE 494048	SITE 512786	SITE 560623	SITE 575415	SITE 598430	SITE 602591	SITE 689326	SITE 696337	SITE 714415	SITE 715532	SITE 846594	SITE 997056	SITE 998666
Study phase	0.32	<0.001	0.01	<0.001	0.002	0.006	0.41	<0.001	0.42	0.02	0.08	0.001	0.07	0.38	0.34	0.33
IRB of record	<0.001	<0.001	*	*	<0.001	0.30	0.005	0.005	0.16	*	*	0.38	*	*	*	0.01
Sponsor type	<0.001	<0.001	<0.001	<0.001	<0.001	0.22	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Pediatric subjects	0.45	0.12	<0.001	0.54	0.91	0.40	<0.001	0.008	0.69	0.003	0.01	0.11	0.002	0.04	0.69	0.99
Randomized	0.55	0.11	0.39	0.32	0.86	0.29	0.02	0.09	0.36	0.005	0.003	0.10	0.27	0.63	0.02	0.79
Placebo	0.001	0.92	0.64	0.06	0.01	0.61	0.93	0.46	0.69	0.22	0.30	0.009	0.87	0.13	0.16	0.87
Primary endpoint	0.42	<0.001	0.18	0.12	0.76	0.01	0.61	0.55	0.76	0.68	0.02	0.08	0.25	0.07	0.23	0.83
National enrollment goal [¥]	0.07	0.30	<0.001	0.01	<0.001	0.01	0.48	0.06	0.04	<0.001	0.37	<0.001	0.98	0.31	0.91	0.10
# Sites nationally	<0.001	<0.001	<0.001	<0.001	<0.001	0.01	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.001
Total # months accruing [¥]	<0.001	<0.001	<0.001	<0.001	<0.001	0.008	<0.001	<0.001	<0.001	<0.001	0.14	<0.001	0.002	<0.001	0.002	<0.001
# Months open prior to local open [¥]	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.17	<0.001	<0.001	<0.001	0.02	0.02
Treatment study	<0.001	<0.001	0.38	<0.001	**	0.03	0.81	<0.001	0.69	0.15	0.05	<0.001	0.72	0.01	0.63	0.85
Precision medicine	0.55	0.16	0.10	0.11	**	0.051	0.45	0.09	0.14	0.46	0.16	0.26	0.18	0.02	0.43	0.84

Disease Team	0.04	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.23	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.18	<0.001
--------------	-------------	------------------	------------------	------------------	------------------	------------------	------------------	------	------------------	------------------	------------------	------------------	------------------	------------------	------	------------------

P-values of the reduced negative binomial regression models of the associated variable with protocol workload (protocols per staff member or FTE) and local accrual. Statistically significant values ($p < 0.05$) are bolded.

** Collinear due to the institution only utilizing one IRB*

*** Data not provided*

¥ Transformed variable used (log transformation for national enrollment and total number of months accruing; square root transformation for number of months of national accrual completed).

4.6.2 Effect Modifiers and Interactions

Four variable pairings were tested for effect modification and/or interaction with local accrual. Each variable pairing was with national enrollment goal and included the variables number of national sites, months of accrual already completed, total months of national accrual, and sponsor type. The summary of the results of these interaction pairings on the outcome of local accrual is in **Table 4-40**. Full models testing interactions as well as without the interaction (including all combinations of interactions for sites with multiple interaction pairings) were tested for best fit. In most cases, a model without any interactions had the best fit. This was observed in sites 104647, 173472, 448155, 512786, 575415, 598430, 602591, 689326, 714415, and site 846594. The full model for site 715532 maintained one interaction, the national enrollment goal with the sponsor type.

Table 4-40: Effect Modification and Interaction Analyses with Clinical Trial Accrual by Site

INTERACTION PAIR WITH NATIONAL ENROLLMENT	# SITES	MONTHS NATIONAL ACCRUAL COMPLETED	TOTAL MONTHS PLANNED ACCRUAL	SPONSOR TYPE
SITE 104647	Effect Modifier	Effect Modifier	Interaction	Interaction
SITE 173472	Effect Modifier	Effect Modifier	Effect Modifier	Interaction
SITE 448155	Effect Modifier	Interaction	Effect Modifier	Interaction
SITE 494048	Effect Modifier	Effect Modifier	Effect Modifier	Effect Modifier

SITE 512786	Interaction	Effect Modifier	Interaction	Interaction
SITE 560623	Effect Modifier	Effect Modifier	Effect Modifier	Effect Modifier
SITE 575415	Effect Modifier	Interaction	Interaction	Effect Modifier
SITE 598430	Effect Modifier	Effect Modifier	Neither	Interaction
SITE 602591	Effect Modifier	Effect Modifier	Interaction	Effect Modifier
SITE 689326	Interaction	Interaction	Effect Modifier	Interaction
SITE 696337	Effect Modifier	Effect Modifier	Effect Modifier	Effect Modifier
SITE 714415	Interaction	Interaction	Effect Modifier	Interaction
SITE 715532	Effect Modifier	Interaction	Effect Modifier	Interaction
SITE 846594	Interaction	Effect Modifier	Effect Modifier	Interaction
SITE 997056	Effect Modifier	Effect Modifier	Effect Modifier	Effect Modifier
SITE 998666	Effect Modifier	Effect Modifier	Effect Modifier	Effect Modifier

Tests with an interaction term that had a corresponding p-value <0.05 defined as “Interaction.” If no interaction, models with a change in the beta coefficient of protocol workload $\geq 10\%$ were defined as “Effect Modifier,” otherwise listed as “Neither.”

4.6.3 Full Prediction Model Results – Site-Specific

Clinical trials used to create the models were run through the site-specific models to assess the site-predicted clinical trial accrual. Each site-specific full model is documented in **Appendix G-1**. All site models were statistically significant at the $p < 0.001$ level. Combined, the site-specific models predicted a total of 55,613 subjects, 112.8% of actual accrual (**Table 4-41**). At the site level, site-specific models predicted between 96% and 207% of actual accrual. The institution

that had 207% of predicted accrual had one outlier that drove the calculation (the prediction model predicted over 3,000 subject accruals). Removing this outlier lowered the predicted accrual to 92% of actual. The institution with the next highest percent of actual predicted was 120%. At the study level, and for studies that accrued at least one subject, the site predicted models predicted, on median, actual accrual at 102% of actual (range 8.3%-2750%; interquartile range 62.9%-189%).

Table 4-41: Actual and Predicted Accrual Models by Site

	ACTUAL ACCRUAL	DISEASE TEAM PREDICTION	SITE-SPECIFIC MODEL	OVERALL MODEL: UNADJUSTED	OVERALL MODEL: ADJUSTED
ALL	49319	109076 (221%)	55613 (113%)	52985 (107%)	52994 (106%)
SITE 104647	3497	7066 (202%)	3530 (101%)	3354 (96%)	3451 (99%)
SITE 173472	3407	5886 (172%)	3599 (106%)	3829 (112%)	3962 (116%)
SITE 448155	6716	9657 (144%)	6717 (100%)	5409 (81%)	5615 (84%)
SITE 494048	4029	8694 (216%)	4073 (101%)	4553 (113%)	4726 (117%)
SITE 512786	6228	12696 (204%)	5955 (96%)	5555 (89%)	5690 (91%)
SITE 560623	363	1206 (332%)	389 (107%)	1151 (317%)	388 (107%)
SITE 575415	3549	6094 (172%)	3531 (99%)	3738 (105%)	3891 (110%)
SITE 598430	1338	7156 (535%)	1321 (99%)	2375 (178%)	2495 (186%)
SITE 602591	2113	6299 (298%)	2177 (103%)	2644 (125%)	2762 (131%)
SITE 689326	6705	18834 (281%)	6470 (96%)	4421 (66%)	4582 (68%)
SITE 696337	831	1917 (231%)	864 (104%)	960 (116%)	931 (112%)
SITE 714415	5530	8629 (156%)	11449 (207%)*	8697 (157%)**	8943 (162%)***
SITE 715532	1313	4756 (362%)	1409 (107%)	1903 (145%)	1218 (93%)
SITE 846594	1854	6586 (355%)	2223 (120%)	2480 (134%)	2562 (138%)
SITE 997056	754	1490 (198%)	752 (100%)	697 (92%)	724 (96%)
SITE 998666	1092	2110 (193%)	1154 (106%)	1219 (112%)	1054 (97%)

*Removing one outlier lowers the number of subjects to *5065 subjects (92% of actual); **4745 subjects (86% of actual);*

****4907 subjects (89% of actual).*

The correlation between site prediction model values and actual values was weak ($\rho=0.1388$) with site-specific correlations ranging from 0.1259-0.8001 (median 0.6802; **Table 4-42**).

Table 4-42: Correlation of Predicted Accrual to Actual Accrual

SITE	DISEASE TEAM PREDICTION	SITE-SPECIFIC MODEL	OVERALL MODEL: UNADJUSTED	OVERALL MODEL: ADJUSTED
ALL	0.4455	0.1388	0.1608	0.1599
104647	0.3128	0.6833	0.4176	0.4191
173472	0.5997	0.5309	0.5276	0.5354
448155	0.6837	0.7676	0.6971	0.7039
494048	0.5715	0.7419	0.4824	0.4859
512786	0.7701	0.6771	0.5414	0.5431
560623	0.4095	0.6088	0.3883	0.3946
575415	0.8546	0.7628	0.7478	0.7447
598430	0.2497	0.6836	0.5984	0.6028
602591	0.4069	0.6695	0.6532	0.6570
689326	0.3526	0.7671	0.7441	0.7403
696337	0.5117	0.5031	0.1560	0.1439
714415	0.9188	0.1259	0.1453	0.1449
715532	0.4255	0.8001	0.6565	0.6511
846594	0.4443	0.4786	0.4951	0.4954
997056	0.8461	0.6544	0.5905	0.5941
998666	0.2879	0.7061	0.7216	0.7258

Pearson correlation coefficient between prediction model and actual accrual.

Removing the one outlier observation, where the site prediction model predicted over 3,000 enrollments, the overall correlation coefficient increased to 0.6903, a moderate correlation (Figure 4-10). These results were categorized by amount of accrual (<1, 1<4, 4<7, 7<10, 10<20, 20<50, and 50+ subjects) and tabulated against actual accrual utilizing the same categories. Each site’s tabulation can be found in Appendix G-2. Overall, sites matched on categories 38.1% of the time using the site-specific model (Table 4-43), with a range of 32.0%-57.5% (median 39.4%; Table 4-44). The comparison group, disease team prediction, was 21.7% overall (Table 4-45) with a range of 5.2%-31.9% (median 19.7%; Table 4-44).

Table 4-43: Site-Specific Prediction Model versus Actual Accrual – All Sites

Actual Accrual	Site Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	252	547	152	52	37	11	2	1053
1 < 4	128	939	428	135	129	27	1	1787
4 < 7	10	348	320	133	116	18	1	946
7 < 10	0	117	183	118	102	29	4	553
10 < 20	0	51	193	166	308	90	3	811
20 < 50	0	2	29	46	203	219	21	520
> 50	0	0	0	4	12	55	46	117
Total	390	2004	1305	654	907	449	78	5787

Table 4-44: Correlation of Categories between Predicted Accrual and Actual Accrual

SITE	DISEASE TEAM PREDICTION	SITE- SPECIFIC MODEL	OVERALL MODEL: UNADJUSTED	OVERALL MODEL: ADJUSTED
ALL	21.7%	38.1%	31.6%	32.7%
104647	29.9%	33.2%	32.3%	34.5%
173472	30.5%	40.0%	34.2%	34.4%
448155	31.9%	36.2%	31.0%	31.0%
494048	19.2%	34.2%	29.5%	31.0%
512786	22.9%	32.0%	30.8%	31.0%
560623	12.3%	57.5%	16.8%	52.5%
575415	27.6%	34.4%	32.7%	32.9%
598430	5.2%	40.0%	28.8%	26.4%
602591	17.0%	41.0%	35.2%	33.2%
689326	19.7%	35.5%	37.1%	36.2%
696337	8.2%	40.7%	24.7%	22.0%
714415	29.3%	40.3%	35.0%	34.8%
715532	10.2%	48.4%	26.2%	34.7%
846594	8.2%	42.9%	32.0%	30.3%
997056	20.4%	38.7%	36.6%	40.9%
998666	19.7%	37.3%	30.6%	27.5%

Number of matching categories between predicted accrual and actual accrual.

Table 4-45: Disease Team Predicted versus Actual Accrual – All Sites

Actual Accrual	Disease Team Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	2	190	356	71	282	94	58	1053
1 < 4	6	215	567	146	593	202	58	1787
4 < 7	0	27	233	77	417	164	28	946
7 < 10	0	8	55	39	311	118	22	553
10 < 20	0	5	51	20	331	335	69	811
20 < 50	0	1	11	5	61	331	111	520
> 50	1	0	0	0	4	8	104	117
Total	9	446	1273	358	1999	1252	450	5787

4.6.4 Full Prediction Model Results – All Sites

Two methods for variable selection for the all-sites models occurred. The first followed the methodology for the site-specific models. Univariate analyses were performed, both the unadjusted and adjusted (using institution as the random effect) results of variable association with local accrual are summarized in **Table 4-46**.

Table 4-46: Independent Variable Association with Clinical Trial Accrual for All Sites

Variable	NEGATIVE BINOMIAL REGRESSION: UNADJUSTED	RANDOM EFFECTS MODEL (EFFECT: INSTITUTION)
Study phase	<0.001	<0.001
IRB of record	<0.001	<0.001
Sponsor type	<0.001	<0.001
Pediatric subjects	<0.001	<0.001
Randomized	<0.001	<0.001
Placebo	<0.001	<0.001
Primary endpoint	0.36	0.30
National enrollment goal $\sqrt{\text{}}$	<0.001	<0.001
# Sites nationally	<0.001	<0.001
Total # months accruing $\sqrt{\text{}}$	<0.001	<0.001
# Months open prior to local open $\sqrt{\text{}}$	<0.001	<0.001

P-values of the unadjusted and adjusted univariate negative binomial regression of the associated variable with local accrual. Statistically significant values ($p < 0.05$) are bolded. $\sqrt{\text{}}$ Transformed variable used (log transformation for national enrollment and total number of months accruing; square root transformation for number of months of national accrual completed).

The second method involved utilizing the variables found most consistently associated with clinical trial accrual at the site level. These variables were sponsor type, national enrollment goal, total months of accrual, months of national enrollment already completed, and the number of sites participating. Both methods of variable selection included testing for effect modification

and interaction on the pre-determined pairings, which found interactions with national enrollment goal and each of the following variables: number of national sites, total number of months of national enrollment, number of months of national enrollment already completed, and sponsor type (Table 4-47).

Table 4-47: Effect Modification and Interaction Analyses with Clinical Trial Accrual for All Sites

INTERACTION PAIR WITH NATIONAL ENROLLMENT	# SITES	MONTHS NATIONAL ACCRUAL COMPLETED	TOTAL MONTHS PLANNED ACCRUAL	SPONSOR TYPE
NEGATIVE BINOMIAL REGRESSION: UNADJUSTED	Interaction	Interaction	Interaction	Interaction
RANDOM EFFECTS MODEL BY INSTITUTION	Interaction	Interaction	Interaction	Interaction

Tests with an interaction term that had a corresponding p-value <0.05 defined as “Interaction.” If no interaction, models with a change in the beta coefficient of either variable $\geq 10\%$ were defined as “Effect Modifier,” otherwise listed as “Neither.”

In both methods of variable selection, there was not a statistically significant difference between the unadjusted and adjusted models (p=1.000). Better model fit, demonstrated by lower BIC values, was seen in the second variable selection method of utilizing the variables correlating with accrual most consistently in the site-specific models. These terms plus the following

interaction terms with national enrollment goal made up the full overall model for all sites: number of national sites, total months of accrual, and sponsor type. Predicted accrual from the unadjusted model and the adjusted model (using institution as the random effect) were compared as the unadjusted model was not statistically different from the adjusted model and the adjusted model provided greater information for accrual prediction. The unadjusted model can be viewed in **Table 4-48** and the adjusted model in **Table 4-49**.

Table 4-48: Full, Unadjusted Model Predicting Clinical Trial Accrual for All Sites

```
Negative binomial regression      Number of obs   =      5,787
LR chi2(18)                      =      2979.50
Dispersion = mean                Prob > chi2     =      0.0000
Log likelihood = -16535.015      Pseudo R2      =      0.0827
```

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

sponsor_type						
Institutitonal	-.1979086	.2187996	-0.90	0.366	-.6267478 .2309307	
Industry	.0114266	.2091969	0.05	0.956	-.3985919 .421445	
National Group	-1.473999	.2320511	-6.35	0.000	-1.928811 -1.019187	
natenroll_trans	.9006595	.0749264	12.02	0.000	.7538064 1.047513	
totalmo_trans	.693492	.0820163	8.46	0.000	.5327431 .8542409	
moaccrdone_trans	-.2972761	.0101305	-29.34	0.000	-.3171316 -.2774206	
natsitescat						
10-49 sites	-.4183979	.1942489	-2.15	0.031	-.7991188 -.037677	
50-199 sites	.010354	.2263233	0.05	0.964	-.4332316 .4539395	
200+ sites	.961147	.2672874	3.60	0.000	.4372733 1.485021	
Unknown	.0555887	.2879496	0.19	0.847	-.5087823 .6199596	
natenroll_trans	0	(omitted)				
natsitescat#c.natenroll_trans						
10-49 sites	-.0474634	.0458861	-1.03	0.301	-.1373985 .0424717	
50-199 sites	-.2190291	.0468915	-4.67	0.000	-.3109348 -.1271234	
200+ sites	-.3711943	.0506578	-7.33	0.000	-.4704818 -.2719068	
Unknown	-.140411	.0605826	-2.32	0.020	-.2591507 -.0216714	
natenroll_trans	0	(omitted)				
totalmo_trans	0	(omitted)				
c.natenroll_trans#c.totalmo_trans	-.073825	.0170367	-4.33	0.000	-.1072164 -.0404337	
natenroll_trans	0	(omitted)				
sponsor_type#c.natenroll_trans						
Institutitonal	.0436578	.0574436	0.76	0.447	-.0689296 .1562452	
Industry	-.1852257	.0519634	-3.56	0.000	-.2870722 -.0833792	
National Group	.0739356	.0543429	1.36	0.174	-.0325747 .1804458	

```

      _cons | -1.459888   .3306492   -4.42   0.000   -2.107948   -.8118271
-----+-----
      /lnalpha | -.0162747   .023375                -.062089   .0295395
-----+-----
      alpha | .983857   .0229977                .9397993   1.02998
-----+-----
LR test of alpha=0: chibar2(01) = 3.1e+04          Prob >= chibar2 = 0.000

```

Measures of Fit for nbreg of local_accrual

```

Log-Lik Intercept Only:  -18024.763   Log-Lik Full Model:      -16535.015
D(5759):                 33070.030   LR(18):                 2979.496
                          Prob > LR:                0.000
McFadden's R2:          0.083   McFadden's Adj R2:      0.081
Maximum Likelihood R2:  0.402   Cragg & Uhler's R2:    0.403
AIC:                    5.724   AIC*n:                 33126.030
BIC:                   -16822.314  BIC':                  -2823.555

```

Table 4-49: Full, Adjusted Random Effects Model Predicting Clinical Trial Accrual for All Sites

```

Mixed-effects nbinoimial regression      Number of obs   =      5,787
Overdispersion:      mean
Group variable:      siteid              Number of groups =      16

Obs per group:
      min =      93
      avg =     361.7
      max =     697

Integration method:  mvaghermite         Integration pts. =      7

Log likelihood = -16575.206              Wald chi2(18)   =     3240.09
                                          Prob > chi2     =      0.0000

```

```

-----+-----
      local_accrual |      Coef.   Std. Err.   z   P>|z|   [95% Conf. Interval]
-----+-----
      |
      sponsor_type |
      Institutional |  -.285857   .2173978   -1.31  0.189   -.7119488   .1402348
      Industry      |  -.0571687   .2083548   -0.27  0.784   -.4655366   .3511992
      National Group | -1.588053   .2312533   -6.87  0.000   -2.041301   -1.134805
      |
      natenroll_trans |  .8795506   .0744883   11.81  0.000   .7335563   1.025545

```

totalmo_trans		.6918171	.0812915	8.51	0.000	.5324886	.8511456
moaccrdone_trans		-.2907135	.0100881	-28.82	0.000	-.3104859	-.270941
natsitescat							
10-49 sites		-.4519167	.1924443	-2.35	0.019	-.8291005	-.0747329
50-199 sites		.0520915	.2248467	0.23	0.817	-.3886	.4927831
200+ sites		.9822869	.2656227	3.70	0.000	.4616759	1.502898
Unknown		.0699725	.2850361	0.25	0.806	-.4886879	.628633
natsitescat#c.natenroll_trans							
10-49 sites		-.0418616	.045455	-0.92	0.357	-.1309517	.0472285
50-199 sites		-.2243699	.0465405	-4.82	0.000	-.3155876	-.1331523
200+ sites		-.3714638	.0502659	-7.39	0.000	-.4699831	-.2729445
Unknown		-.1449464	.0599945	-2.42	0.016	-.2625334	-.0273594
c.natenroll_trans#c.totalmo_trans		-.0729931	.016895	-4.32	0.000	-.1061067	-.0398794
sponsor_type#c.natenroll_trans							
Instituitonal		.0624491	.0569707	1.10	0.273	-.0492114	.1741096
Industry		-.1674306	.0517153	-3.24	0.001	-.2687907	-.0660706
National Group		.1037317	.0541098	1.92	0.055	-.0023215	.2097849
_cons		-1.363591	.3287852	-4.15	0.000	-2.007998	-.7191841

/lnalpha		-.0492488	.023558	-2.09	0.037	-.0954216	-.003076

_all>siteid							
var(_cons)		.3304592	.2148674			.0923973	1.181888

LR test versus nbinomial model: chibar2(01) = 0.00 Prob >= chibar2 = 1.0000

Akaike's information criterion and Bayesian information criterion

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	5,787	.	-16575.21	21	33192.41	33332.34

Combined, both the unadjusted and adjusted overall models predicted approximately 52,990 subjects, 107.4% of actual accrual (**Table 4-41**; page 240). At the site level, the unadjusted overall model predicted between 66% and 178% of actual accrual for the unadjusted model and 68%-186% for the adjusted model. At the study level, for studies that accrued at least one subject, both overall models predicted, on median, accrual at 112% of actual. For the unadjusted model, the range was 3.0%-6,174%; interquartile range of 65.0%-223%. The range of the adjusted model was 3.1%-6,306%; interquartile range of 65.1%-221.0%.

The correlation between each overall prediction model values and actual values was weak ($\rho=0.1608$, unadjusted and $\rho=0.1599$, adjusted) with site-specific correlations ranging from 0.1453-0.7478 (median 0.5945, unadjusted) and 0.1439-0.7447 (median 0.5985, adjusted; **Table 4-42**; page 241). Correlation values were driven by one outlier where all prediction regression models predicted >3,000 subjects. Removing this outlier, the correlation coefficient increased to ~0.68, a moderate correlation. Results were categorized by amount of accrual (<1, 1<4, 4<7, 7<10, 10<20, 20<50, and 50+ subjects) and tabulated against actual accrual utilizing the same categories. The overall tabulation can be found in **Table 4-50** (unadjusted) and **Table 4-51** (adjusted). Each site's tabulation can be found in **Appendix G-2** through **Appendix G-5**.

Table 4-50: Overall Prediction Model versus Actual Accrual – All Sites

Actual Accrual	Overall Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	65	532	280	92	66	11	7	1053
1 < 4	34	778	613	198	135	27	2	1787
4 < 7	3	299	371	155	98	17	3	946
7 < 10	0	101	201	126	98	27	0	553
10 < 20	1	75	223	181	257	72	2	811
20 < 50	0	13	60	61	181	197	8	520
> 50	0	1	2	6	22	53	33	117
Total	103	1799	1750	819	857	404	55	5787

Table 4-51: Overall Prediction Model: Random Effects versus Actual Accrual – All Sites

Actual Accrual	Overall Model Predicted Accrual: Random Effects							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	108	519	260	76	72	11	7	1053
1 < 4	45	779	596	207	136	22	2	1787
4 < 7	3	286	379	146	110	20	2	946
7 < 10	0	100	203	121	100	29	0	553
10 < 20	1	72	218	174	273	72	1	811
20 < 50	0	9	60	56	185	198	9	520
> 50	0	1	3	3	22	55	33	117
Total	157	1766	1719	786	898	407	54	5787

Overall, sites matched on categories 31.6% (unadjusted) and 32.7% (adjusted) of the time using the site-specific model. At the site level, the unadjusted model matched sites 16.8%-37.1% of

the time (median 31.5%), while the adjusted model had a range of 22.0%-52.5% (median 33.1%). Specific site values are listed in **Table 4-44** (page 243). The comparison group, disease team prediction, was 21.7% overall with a range of 5.2%-31.9% (median 19.7%).

4.6.5 Model Sensitivity, Specificity, and Accuracy

Defining four accruals as a cut-off value for “success,” two-way tables were constructed to assess how well the disease team, site-specific models, unadjusted model, and adjusted model predicted local accrual. Three measures were assessed: sensitivity, specificity, and model accuracy. The values for all sites combined as well as each site individually are listed in **Table 4-52**. Overall, the disease teams were 57.3% accurate on determining whether or not a clinical trial would accrue at least four subjects. Teams were 98.6% sensitive and 14.5% specific. At the site level, disease teams ranged from 28.5%-71.5% accurate (median 54.7%). No team was under 95% sensitive in predicting accrual (median 98.8%) or over 30% specific (range 1.5%-29.6%; median 16.0%).

Table 4-52: Sensitivity, Specificity, and Accuracy for Accrual Prediction Models using a Four Accrual Cut-off

SITE	DISEASE TEAM PREDICTED ACCRUAL			SITE ACCRUAL PREDICTION MODEL			OVERALL ACCRUAL PREDICTION MODEL			OVERALL ACCRUAL PREDICTION MODEL: RANDOM EFFECTS		
	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy
ALL	98.6%	14.5%	57.3%	82.1%	65.7%	74.0%	83.3%	49.6%	66.8%	84.0%	51.1%	67.8%
104647	95.8%	19.7%	64.1%	86.4%	46.7%	69.9%	80.3%	46.1%	66.0%	85.9%	42.8%	67.9%
173472	97.0%	29.6%	62.9%	82.1%	68.9%	75.4%	87.1%	55.8%	71.3%	89.6%	53.4%	71.2%
448155	97.3%	24.9%	66.9%	86.6%	59.4%	75.2%	81.4%	53.2%	69.6%	82.7%	50.5%	69.2%
494048	98.8%	16.7%	53.4%	77.7%	67.5%	72.1%	87.4%	43.6%	63.2%	91.5%	41.6%	63.9%
512786	100.0%	4.4%	68.8%	92.9%	24.1%	70.5%	85.0%	35.4%	68.8%	86.5%	32.9%	69.0%
560623	96.6%	15.3%	28.5%	51.7%	90.7%	84.4%	86.2%	48.7%	54.7%	17.2%	92.7%	80.4%
575415	99.5%	16.9%	60.2%	79.5%	63.6%	72.0%	83.3%	48.7%	66.8%	85.6%	46.1%	66.8%
598430	100.0%	3.3%	36.5%	58.9%	82.2%	74.2%	82.1%	50.0%	61.0%	83.9%	44.4%	58.0%
602591	100.0%	2.7%	45.8%	70.9%	77.3%	74.4%	82.9%	55.5%	67.6%	85.1%	50.0%	65.6%
689326	100.0%	1.5%	71.5%	95.4%	32.6%	77.2%	83.3%	53.8%	74.8%	86.7%	48.5%	75.7%
696337	100.0%	7.6%	46.7%	72.7%	81.9%	78.0%	71.4%	46.7%	57.1%	67.5%	48.6%	56.6%
714415	98.0%	19.0%	63.3%	81.5%	59.9%	72.0%	83.5%	54.3%	70.7%	85.2%	50.4%	69.9%
715532	98.8%	18.1%	47.6%	70.7%	81.8%	77.8%	86.6%	43.4%	59.1%	69.5%	72.0%	71.1%
846594	100.0%	10.1%	45.2%	76.5%	74.3%	75.2%	85.2%	53.6%	66.0%	85.2%	50.3%	63.9%
997056	100.0%	4.7%	55.9%	86.0%	55.8%	72.0%	84.0%	48.8%	67.7%	86.0%	46.5%	67.7%
998666	95.0%	19.5%	50.8%	66.3%	77.0%	72.5%	76.3%	50.4%	61.1%	67.5%	61.9%	64.2%

The site-specific models were more accurate in predicting whether or not a protocol would accrue at least four subjects than the disease team in all 16 site-specific models. Overall, the site-specific models were 74.0% accurate. Accuracy of the model for each site ranged from 69.9%-84.4% (median 74.3%). The absolute improvement in the site-specific model accuracy from disease team accuracy ranged from 1.7%-55.9% (median 17.4%). Site-specific model sensitivity ranged from 51.7%-92.9% (median 78.8%); specificity ranged from 24.1%-90.7% (median 68.2%). Using a cut-off value of four subjects, the number of studies that the model classified beneath this value that actually accrued more than four subjects (false negatives) was 528 (9.1% of total, 22% of studies predicted to accrue fewer than four subjects). If the model was solely used to determine whether a study should be opened then, based on a cut-off value of four subjects, 22% of studies chosen not to be opened would have been done so incorrectly. Conversely, the model predicted 974 studies would accrue at least four subjects that actually did not meet that benchmark (false positives). This was 34% of the number of studies that did not accrue at least four subjects.

The unadjusted overall model predicted accrual more accurately than the disease teams at 15 sites and predicted the same accuracy at one site (68.8%). Overall, the unadjusted model was 66.8% accurate, less than the site-specific models. Site accuracy using the unadjusted model ranged from 54.7%-74.8% (median 66.4%). The site-specific models were more accurate than the unadjusted model at all 16 sites. Unadjusted model sensitivity was 83.3% combining all sites, with a range of 71.4%-87.4% at the site level (median 83.4%). Specificity for the unadjusted model was 49.6% with a range of 35.4%-55.8% (median 49.4%).

The adjusted overall model predicted accrual more accurately than the disease teams at all 16 sites. Overall, at all sites, the adjusted model was 67.8% accurate, less accurate than the site-specific models but one percentage point more specific than the unadjusted model. The range of accuracy at sites was 56.6%-80.4% (median 67.8%). Sensitivity of the adjusted model was 84.0% overall, with a range of 17.2%-91.5% and a median of 85.2%. The second lowest sensitivity was 67.5%. Model specificity was 51.1% overall with a range of 32.9%-92.7% (median 49.3%).

Using a cut-off value of four subjects, the number of studies that the unadjusted model classified beneath this value that actually accrued more was 493 (8.5% of total, 25.9% of studies predicted to accrue fewer than four subjects). If the model was solely used to determine whether a study should be opened then, based on a cut-off value of four subjects, 25.9% of studies chosen not to be opened would have been done so incorrectly. Conversely, the unadjusted model predicted 1,431 studies would accrue at least four subjects that actually did not meet that benchmark. This was 50.4% of the number of studies that did not accrue at least four subjects. For the adjusted model, the number of studies that accrued at least four subjects but predicted not to was slightly lower, at 475 (8.2% of total, 24.7% of studies predicted to accrue fewer than four subjects). The adjusted model predicted 1,389 studies that would accrue at least four subjects that failed to accrue four subjects, 48.9% of the number of studies that did not accrue at least four subjects.

4.6.6 Specific Aim Three Summary

- Specific Aim 3a Finding: An overall, adjusted prediction model was generated that described clinical trial accrual for all sites
- Specific Aim 3b Finding: The site-specific models predicted clinical trial accrual better than the overall model at all 16 centers (using a binary cut-off value of four accruals)
- Specific Aim 3b Finding: The site-specific models predicted clinical trial accrual better than the overall model at 13 of 16 centers (matching accrual categories)
- Specific Aim 3c Finding: The site-specific models predicted clinical trial accrual better than the disease teams at all 16 centers (using a binary cut-off value of four accruals)
- Specific Aim 3c Finding: The site-specific models predicted clinical trial accrual better than the disease teams at all 16 centers (matching accrual categories).

Overall, the site-specific models predicted 55,613 subjects, or 113% of actual and the overall models predicted approximately 52,990 subjects, or 107% of actual. In comparison, disease teams, the current method of predicting subject accrual, predicted 109,076 subjects, or 221% of actual. A summary of the best prediction model overall and for each site is presented in **Table 4-53**. In the binary models (predicting whether a study would or would not accrue at least four subjects), the site-specific models were the most accurate models in all 16 cancer centers. The disease team predictions high the highest correlation with actual accrual, according to the Pearson correlation coefficient; however, when categorizing accrual into categories (<1, 1<4, 4<7, 7<10, 10<20, 20<50, 50+), the site-specific model matched the highest percentage of categories overall and in 13 of the 16 sites. For studies that were incorrectly classified by the

prediction models, the model was more likely to favor opening the study (higher percentage of studies that should have accrued at least four subjects and did not).

Table 4-53: Best Accrual Prediction Model by Site

SITE	PEARSON CORRELATION	% CATEGORIES MATCHING	BINARY (<4/4+) MATCHING: ACCURACY
104647	Disease Team	Site Model	Site Model
173472	Site Model	Overall Random Effects Model	Site Model
448155	Disease Team	Site Model	Site Model
494048	Site Model	Site Model	Site Model
512786	Site Model	Site Model	Site Model
560623	Disease Team	Site Model	Site Model
575415	Site Model	Site Model	Site Model
598430	Disease Team	Site Model	Site Model
602591	Site Model	Site Model	Site Model
689326	Site Model	Site Model	Site Model
696337	Site Model	Overall Model	Site Model
714415	Disease Team	Site Model	Site Model
715532	Disease Team	Site Model	Site Model
846594	Site Model	Site Model	Site Model
997056	Overall Random Effects Model	Site Model	Site Model
998666	Disease Team	Overall Random Effects Model	Site Model

4.7 Site-Specific Summaries

4.7.1 Site 104647

The number of protocols submitted was 365, contributing 6.3% to the total number of protocols studied in this project. Of these, 47 (12.9%) accrued zero subjects. The number of protocols that accrued at least four subjects was 213 (58.4%). Total contributed accrual was 3,497 subjects (7.1% of overall accrual). The distribution of protocol accrual is displayed in **Appendix E-1a**.

Activation times for this institution were not provided. This site instated a feasibility committee in 2014 for protocols with a sponsor type of National Group. The site reported a centralized clinical trials office with some groups utilizing the regulatory functions of the office. Staff are not dedicated solely to regulatory functions, therefore staff workload did differ between protocols per staff member and protocols per FTE.

4.7.1.1 Accrual Prediction

Disease teams predicted a total of 7,066 subjects, 202% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the disease team correctly predicted it would 95.8% of the time. Disease teams correctly predicted 19.7% of the time that a study would not accrue four subjects. Total accuracy of the disease team on this binary measure was 64.1% (**Table 4-52**; page 254). Disease teams predicted the category of actual accrual 29.9% of the time (**Appendix G-3a**).

The site-specific prediction model predicted a total of 3,530 subjects, 101% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the site-specific prediction model correctly predicted that studies would accrue at least four subjects 86.4% of the time. The model correctly predicted 46.7% of the time that a study would not accrue four subjects. Total accuracy of the site-specific model was 69.9% (**Table 4-52**; page 254). The site-specific model predicted the category of actual accrual 33.2% of the time (**Appendix G-2a**). When the trial actually accrued at least four subjects, both the disease team and the site-specific model predicted it would 82.6% of the time.

The unadjusted, overall prediction model predicted a total of 3,354 subjects, 96% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 80.3% of the time. The model correctly predicted 46.1% of the time that a study would not accrue four subjects. Total accuracy of the unadjusted, overall model was 66.0% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 32.3% of the time (**Appendix G-4a**).

The adjusted, overall prediction model predicted a total of 3,451 subjects, 99% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 85.9% of the time. The model correctly predicted 42.8% of the time that a study would not accrue four subjects. Total accuracy of the adjusted, overall model was 67.9% (**Table**

4-52; page 254). The unadjusted, overall model predicted the category of actual accrual 34.5% of the time (**Appendix G-5a**).

4.7.2: Site 173472

The number of protocols submitted was 407, contributing 7.0% to the total number of protocols studied in this project. Of these, 71 (17.4%) accrued zero subjects. The number of protocols that accrued at least four subjects was 201 (49.4%). Total contributed accrual was 3,407 subjects (6.9% of overall accrual). The distribution of protocol accrual is displayed in **Appendix E-1b**.

The average time to activate a study was 279.2 ± 232.8 days (median 217 days). This was slower than the overall activation time of 190.1 ± 164.0 days (median 154 days). This site did not utilize a feasibility committee. Information regarding the usage and regulatory functions of the center was not provided.

4.7.2.1 Accrual Prediction

Disease teams predicted a total of 5,886 subjects, 172% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the disease team correctly predicted it would 97.0% of the time. Disease teams correctly predicted 29.6% of the time that a study would not accrue four subjects. Total accuracy of the disease team on this binary measure was 62.9% (**Table 4-52**; page 254). Disease teams predicted the category of actual accrual 30.5% of the time (**Appendix G-3b**).

The site-specific prediction model predicted a total of 3,599 subjects, 106% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the site-specific prediction model correctly predicted that studies would accrue at least four subjects 82.1% of the time. The model correctly predicted 68.9% of the time that a study would not accrue four subjects. Total accuracy of the site-specific model was 75.4% (**Table 4-52**; page 254). The site-specific model predicted the category of actual accrual 40.0% of the time (**Appendix G-2b**). When the trial actually accrued at least four subjects, both the disease team and the site-specific model predicted it would 81.1% of the time.

The unadjusted, overall prediction model predicted a total of 3,829 subjects, 112% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 87.1% of the time. The model correctly predicted 55.8% of the time that a study would not accrue four subjects. Total accuracy of the unadjusted, overall model was 71.3% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 34.2% of the time (**Appendix G-4b**).

The adjusted, overall prediction model predicted a total of 3,962 subjects, 116% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 89.6% of the time. The model correctly predicted 53.4% of the time that a study would not accrue four subjects. Total accuracy of the adjusted, overall model was 71.2% (**Table**

4-52; page 254). The unadjusted, overall model predicted the category of actual accrual 34.4% of the time (**Appendix G-5b**).

4.7.3: Site 448155

The number of protocols submitted was 697, contributing 12.0% to the total number of protocols studied in this project. Of these, 100 (14.4%) accrued zero subjects. The number of protocols that accrued at least four subjects was 404 (58.0%). Total contributed accrual was 6,716 subjects (13.6% of overall accrual). The distribution of protocol accrual is displayed in **Appendix E-1c**.

The average time to activate a study was 168.7 ± 135.2 days (median 138.5 days). This was faster than the overall activation time of 190.1 ± 164.0 days (median 154 days). This site did not utilize a feasibility committee. The site reported a centralized clinical trials office with some groups utilizing the regulatory functions conducted through the office. Staff are dedicated solely to regulatory functions; however, due to part-time staff, staff workload did differ between protocols per staff member and protocols per FTE. The distribution of protocol workload by staffing model and outcomes are diagrammed in **Appendix F-1a** through **F-6a**.

4.7.3.1 Protocol Workload

For site 448155, the optimal protocol workloads were:

- To maximize clinical trial accrual
 - $2.5 < 3.0$ protocols per staff member by study characteristic analysis (**Appendix F-7a**)

- 2.5 < 5.0 protocols per FTE by study characteristic analysis (**Appendix F-7b**)
- No significant association per staff member/FTE by regression modeling (**Table 4-32**)
- To minimize the percentage of zero-accruing protocols
 - 2.5 < 3.0 protocols per staff member by study characteristic analysis (**Appendix F-7a**)
 - 3.5 < 4.0 protocols per FTE by study characteristic analysis (**Appendix F-7b**)
 - No significant association per staff member/FTE by regression modeling (**Table 4-32**)
- To maximize the percentage of protocols accruing at least four subjects
 - 3.5 < 4.0 protocols per staff member/FTE by study characteristic analysis (**Appendix F-7a** and **Appendix F-7b**)
 - No significant association per staff member/FTE by regression modeling (**Table 4-32**)
- To minimize time to activate a protocol
 - 3.0 < 3.5 protocols per staff member by study characteristic analysis (**Appendix F-7a**)
 - 1.5 < 2.0 protocols per FTE by study characteristic analysis (**Appendix F-7b**)
 - No significant association per staff member/FTE by regression modeling (**Table 4-32**)

4.7.3.2 Accrual Prediction

Disease teams predicted a total of 9,657 subjects, 144% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the disease team correctly predicted it would 97.3% of the time. Disease teams correctly predicted 24.9% of the time that a study would not accrue four subjects. Total accuracy of the disease team on this binary measure was 66.9% (**Table 4-52**; page 254). Disease teams predicted the category of actual accrual 31.9% of the time (**Appendix G-3c**).

The site-specific prediction model predicted a total of 6,717 subjects, 100% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the site-specific prediction model correctly predicted that studies would accrue at least four subjects 86.6% of the time. The model correctly predicted 59.4% of the time that a study would not accrue four subjects. Total accuracy of the site-specific model was 75.2% (**Table 4-52**; page 254). The site-specific model predicted the category of actual accrual 36.2% of the time (**Appendix G-2c**). When the trial actually accrued at least four subjects, both the disease team and the site-specific model predicted it would 85.4% of the time.

The unadjusted, overall prediction model predicted a total of 5,409 subjects, 81% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 81.4% of the time. The model correctly predicted 53.2% of the time that a study would not accrue four subjects. Total accuracy of the unadjusted, overall model was

69.6% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 31.0% of the time (**Appendix G-4c**).

The adjusted, overall prediction model predicted a total of 5,615 subjects, 84% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 82.7% of the time. The model correctly predicted 50.5% of the time that a study would not accrue four subjects. Total accuracy of the adjusted, overall model was 69.2% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 31.0% of the time (**Appendix G-5c**).

4.7.4: Site 494048

The number of protocols submitted was 552, contributing 9.5% to the total number of protocols studied in this project. Of these, 118 (21.4%) accrued zero subjects. The number of protocols that accrued at least four subjects was 247 (44.8%). Total contributed accrual was 4,029 subjects (8.2% of overall accrual). The distribution of protocol accrual is displayed in **Appendix E-1d**.

The average time to activate a study was 124.1 ± 107.9 days (median 100 days). This was faster than the overall activation time of 190.1 ± 164.0 days (median 154 days). This site did utilize a feasibility committee. The site reported a centralized clinical trials office with all regulatory functions conducted through the office. Staff are dedicated solely to regulatory functions, therefore staff workload did not differ between protocols per staff member and protocols per

FTE. The distribution of protocol workload by staffing model and outcomes are diagrammed in **Appendix F-1b** through **F-6b**.

4.7.4.1 Protocol Workload

For site 494048, the optimal protocol workloads were:

- To maximize clinical trial accrual
 - 6.0 < 6.5 protocols per staff member/FTE by study characteristic analysis
(**Appendix F-7c**)
 - 5.5 < 6.0 protocols per staff member/FTE by regression modeling (**Appendix F-8a**)
- To minimize the percentage of zero-accruing protocols
 - 6.0 < 6.5 protocols per staff member/FTE by study characteristic analysis
(**Appendix F-7c**)
 - 6.0 < 6.5 protocols per staff member/FTE by regression modeling (**Appendix F-9a**)
- To maximize the percentage of protocols accruing at least four subjects
 - 3.0 < 3.5 protocols per staff member/FTE by study characteristic analysis
(**Appendix F-7c**)
 - 3.0 < 3.5 protocols per staff member/FTE by regression modeling (**Appendix F-10a**)
- To minimize time to activate a protocol

- < 1 protocol per staff member/FTE by study characteristic analysis (only two protocols reported); 2.5 < 3.0 protocols per staff member/FTE (more than two protocols reported) (**Appendix F-7c**)
- 1.0 < 1.5 protocols per staff member/FTE by regression modeling (**Appendix F-11a**)

4.7.4.2 Accrual Prediction

Disease teams predicted a total of 8,694 subjects, 216% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the disease team correctly predicted it would 98.8% of the time. Disease teams correctly predicted 16.7% of the time that a study would not accrue four subjects. Total accuracy of the disease team on this binary measure was 53.4% (**Table 4-52**; page 254). Disease teams predicted the category of actual accrual 19.2% of the time (**Appendix G-3d**).

The site-specific prediction model predicted a total of 4,073 subjects, 101% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the site-specific prediction model correctly predicted that studies would accrue at least four subjects 77.7% of the time. The model correctly predicted 67.5% of the time that a study would not accrue four subjects. Total accuracy of the site-specific model was 72.1% (**Table 4-52**; page 254). The site-specific model predicted the category of actual accrual 34.2% of the time (**Appendix G-2d**). When the trial actually accrued at least four subjects, both the disease team and the site-specific model predicted it would 77.3% of the time.

The unadjusted, overall prediction model predicted a total of 4,553 subjects, 113% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 87.4% of the time. The model correctly predicted 43.6% of the time that a study would not accrue four subjects. Total accuracy of the unadjusted, overall model was 63.2% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 29.5% of the time (**Appendix G-4d**).

The adjusted, overall prediction model predicted a total of 4,726 subjects, 117% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 91.5% of the time. The model correctly predicted 41.6% of the time that a study would not accrue four subjects. Total accuracy of the adjusted, overall model was 63.9% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 31.0% of the time (**Appendix G-5d**).

4.7.5: Site 512786

The number of protocols submitted was 484, contributing 8.4% to the total number of protocols studied in this project. Of these, 46 (9.5%) accrued zero subjects. The number of protocols that accrued at least four subjects was 326 (67.4%). Total contributed accrual was 6,228 subjects (12.6% of overall accrual). The distribution of protocol accrual is displayed in **Appendix E-1e**.

The average time to activate a study was 227.4 ± 191.5 days (median 181.5 days). This was slower than the overall activation time of 190.1 ± 164.0 days (median 154 days). This site did not utilize a feasibility committee. The site reported a centralized clinical trials office with all regulatory functions conducted through the office. Staff are dedicated solely to regulatory functions, therefore staff workload did not differ between protocols per staff member and protocols per FTE. The distribution of protocol workload by staffing model and outcomes are diagrammed in **Appendix F-1c** through **F-6c**.

4.7.5.1 Protocol Workload

For site 512786, the optimal protocol workloads were:

- To maximize clinical trial accrual
 - $5.0 < 5.5$ protocols per staff member/FTE by study characteristic analysis
(**Appendix F-7d**)
 - No significant association per staff member/FTE by regression modeling (**Table 4-32**)
- To minimize the percentage of zero-accruing protocols
 - $5.0 < 5.5$ protocols per staff member/FTE by study characteristic analysis
(**Appendix F-7d**)
 - No significant association per staff member/FTE by regression modeling (**Table 4-32**)
- To maximize the percentage of protocols accruing at least four subjects

- 2.0 < 2.5 protocols per staff member/FTE by study characteristic analysis
(**Appendix F-7d**)
- 3.5 < 4.0 protocols per staff member/FTE by regression modeling (**Appendix F-10b**)
- To minimize time to activate a protocol
 - 1.5 < 2.0 protocols per staff member/FTE by study characteristic analysis
(**Appendix F-7d**)
 - No significant association per staff member/FTE by regression modeling (**Table 4-32**)

4.7.5.2 Accrual Prediction

Disease teams predicted a total of 12,696 subjects, 204% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the disease team correctly predicted it would 100% of the time. Disease teams correctly predicted 4.4% of the time that a study would not accrue four subjects. Total accuracy of the disease team on this binary measure was 68.8% (**Table 4-52**; page 254). Disease teams predicted the category of actual accrual 22.9% of the time (**Appendix G-3e**).

The site-specific prediction model predicted a total of 5,955 subjects, 96% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the site-specific prediction model correctly predicted that studies would accrue at least four subjects 92.9% of the time. The model correctly predicted 24.1% of the time

that a study would not accrue four subjects. Total accuracy of the site-specific model was 70.5% (**Table 4-52**; page 254). The site-specific model predicted the category of actual accrual 32.0% of the time (**Appendix G-2e**). When the trial actually accrued at least four subjects, both the disease team and the site-specific model predicted it would 92.9% of the time.

The unadjusted, overall prediction model predicted a total of 5,555 subjects, 89% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 85.0% of the time. The model correctly predicted 35.4% of the time that a study would not accrue four subjects. Total accuracy of the unadjusted, overall model was 68.8% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 30.8% of the time (**Appendix G-4e**).

The adjusted, overall prediction model predicted a total of 5,690 subjects, 91% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 86.5% of the time. The model correctly predicted 32.9% of the time that a study would not accrue four subjects. Total accuracy of the adjusted, overall model was 69.0% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 31.0% of the time (**Appendix G-5e**).

4.7.6: Site 560623

The number of protocols submitted was 179, contributing 3.1% to the total number of protocols studied in this project. Of these, 81 (45.3%) accrued zero subjects. The number of protocols that accrued at least four subjects was 29 (16.2%). Total contributed accrual was 363 subjects (0.7% of overall accrual). The distribution of protocol accrual is displayed in **Appendix E-1f**.

The average time to activate a study was 167.8 ± 130.1 days (median 127.5 days). This was faster than the overall activation time of 190.1 ± 164.0 days (median 154 days). This site did not utilize a feasibility committee. The site reported a centralized clinical trials office with all regulatory functions conducted through the office. Staff are dedicated solely to regulatory functions, therefore staff workload did not differ between protocols per staff member and protocols per FTE. The distribution of protocol workload by staffing model and outcomes are diagrammed in **Appendix F-1d** through **F-6d**.

4.7.6.1 Protocol Workload

For site 560623, the optimal protocol workloads were:

- To maximize clinical trial accrual
 - $4.0 < 4.5$ protocols per staff member/FTE by study characteristic analysis
(**Appendix F-7e**)
 - $4.5 < 5.0$ protocols per staff member/FTE by regression modeling (**Appendix F-8b**)
- To minimize the percentage of zero-accruing protocols

- 4.0 < 4.5 protocols per staff member/FTE by study characteristic analysis
(**Appendix F-7e**)
- 4.0 < 4.5 protocols per staff member/FTE by regression modeling (**Appendix F-9b**)
- To maximize the percentage of protocols accruing at least four subjects
 - 4.5 < 5.0 protocols per staff member/FTE by study characteristic analysis
(**Appendix F-7e**)
 - No significant association per staff member/FTE by regression modeling (**Table 4-32**)
- To minimize time to activate a protocol
 - 2.0 < 2.5 protocols per staff member/FTE by study characteristic analysis
(**Appendix F-7e**)
 - No significant association per staff member/FTE by regression modeling (**Table 4-32**)

4.7.6.2 Accrual Prediction

Disease teams predicted a total of 1,206 subjects, 332% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the disease team correctly predicted it would 96.6% of the time. Disease teams correctly predicted 15.3% of the time that a study would not accrue four subjects. Total accuracy of the disease team on this binary measure was 28.5% (**Table 4-52**; page 254). Disease teams predicted the category of actual accrual 12.3% of the time (**Appendix G-3f**).

The site-specific prediction model predicted a total of 389 subjects, 107% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the site-specific prediction model correctly predicted that studies would accrue at least four subjects 51.7% of the time. The model correctly predicted 90.7% of the time that a study would not accrue four subjects. Total accuracy of the site-specific model was 84.4% (**Table 4-52**; page 254). The site-specific model predicted the category of actual accrual 57.5% of the time (**Appendix G-2f**). When the trial actually accrued at least four subjects, both the disease team and the site-specific model predicted it would 51.7% of the time.

The unadjusted, overall prediction model predicted a total of 1,151 subjects, 317% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 86.2% of the time. The model correctly predicted 48.7% of the time that a study would not accrue four subjects. Total accuracy of the unadjusted, overall model was 54.7% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 16.8% of the time (**Appendix G-4f**).

The adjusted, overall prediction model predicted a total of 388 subjects, 107% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 17.2% of the time. The model correctly predicted 92.7% of the time that a study would not accrue four subjects. Total accuracy of the adjusted, overall model was 80.4% (**Table**

4-52; page 254). The unadjusted, overall model predicted the category of actual accrual 52.5% of the time (**Appendix G-5f**).

4.7.7: Site 575415

The number of protocols submitted was 410, contributing 7.1% to the total number of protocols studied in this project. Of these, 83 (20.2%) accrued zero subjects. The number of protocols that accrued at least four subjects was 215 (52.4%). Total contributed accrual was 3549 subjects (7.2% of overall accrual). The distribution of protocol accrual is displayed in **Appendix E-1g**.

Activation times, usage of a feasibility committee, and regulatory functions of staff within and/or external to the centralized clinical trials office were not provided by the site.

4.7.7.1 Accrual Prediction

Disease teams predicted a total of 6,094 subjects, 172% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the disease team correctly predicted it would 99.5% of the time. Disease teams correctly predicted 16.9% of the time that a study would not accrue four subjects. Total accuracy of the disease team on this binary measure was 60.2% (**Table 4-52**; page 254). Disease teams predicted the category of actual accrual 27.6% of the time (**Appendix G-3g**).

The site-specific prediction model predicted a total of 3,531 subjects, 99% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would

be accrued to a study, the site-specific prediction model correctly predicted that studies would accrue at least four subjects 79.5% of the time. The model correctly predicted 63.6% of the time that a study would not accrue four subjects. Total accuracy of the site-specific model was 72.0% (**Table 4-52**; page 254). The site-specific model predicted the category of actual accrual 34.4% of the time (**Appendix G-2g**). When the trial actually accrued at least four subjects, both the disease team and the site-specific model predicted it would 79.5% of the time.

The unadjusted, overall prediction model predicted a total of 3,738 subjects, 105% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 83.3% of the time. The model correctly predicted 48.7% of the time that a study would not accrue four subjects. Total accuracy of the unadjusted, overall model was 66.8% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 32.7% of the time (**Appendix G-4g**).

The adjusted, overall prediction model predicted a total of 3,891 subjects, 110% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 85.6% of the time. The model correctly predicted 46.1% of the time that a study would not accrue four subjects. Total accuracy of the adjusted, overall model was 66.8% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 32.9% of the time (**Appendix G-5g**).

4.7.8: Site 598430

The number of protocols submitted was 326, contributing 5.6% to the total number of protocols studied in this project. Of these, 98 (30.1%) accrued zero subjects. The number of protocols that accrued at least four subjects was 112 (34.4%). Total contributed accrual was 1,338 subjects (2.7% of overall accrual). The distribution of protocol accrual is displayed in **Appendix E-1h**. This site did not utilize a feasibility committee. Activation times and regulatory functions within the center and/or centralized clinical trials office were not provided by the institution.

4.7.8.1 Accrual Prediction

Disease teams predicted a total of 7,156 subjects, 535% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the disease team correctly predicted it would 100% of the time. Disease teams correctly predicted 3.3% of the time that a study would not accrue four subjects. Total accuracy of the disease team on this binary measure was 36.5% (**Table 4-52**; page 254). Disease teams predicted the category of actual accrual 5.2% of the time (**Appendix G-3h**).

The site-specific prediction model predicted a total of 1,321 subjects, 99% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the site-specific prediction model correctly predicted that studies would accrue at least four subjects 58.9% of the time. The model correctly predicted 82.2% of the time that a study would not accrue four subjects. Total accuracy of the site-specific model was 74.2% (**Table 4-52**; page 254). The site-specific model predicted the category of actual accrual 40.0%

of the time (**Appendix G-2h**). When the trial actually accrued at least four subjects, both the disease team and the site-specific model predicted it would 58.9% of the time.

The unadjusted, overall prediction model predicted a total of 2,375 subjects, 178% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 82.1% of the time. The model correctly predicted 50.0% of the time that a study would not accrue four subjects. Total accuracy of the unadjusted, overall model was 61.0% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 28.8% of the time (**Appendix G-4h**).

The adjusted, overall prediction model predicted a total of 2,495 subjects, 186% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 83.9% of the time. The model correctly predicted 44.4% of the time that a study would not accrue four subjects. Total accuracy of the adjusted, overall model was 58.0% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 26.4% of the time (**Appendix G-5h**).

4.7.9: Site 602591

The number of protocols submitted was 395, contributing 6.8% to the total number of protocols studied in this project. Of these, 94 (23.8%) accrued zero subjects. The number of protocols that

accrued at least four subjects was 175 (44.3%). Total contributed accrual was 2,113 subjects (4.3% of overall accrual). The distribution of protocol accrual is displayed in **Appendix E-1i**.

The average time to activate a study was 220.1 ± 131.9 days (median 202 days). This was slower than the overall activation time of 190.1 ± 164.0 days (median 154 days). This site did not utilize a feasibility committee. The site reported a centralized clinical trials office with all regulatory functions conducted through the office. Staff are dedicated solely to regulatory functions; however, due to part-time staff, staff workload did differ between protocols per staff member and protocols per FTE. The distribution of protocol workload by staffing model and outcomes are diagrammed in **Appendix F-1e** through **F-6e**.

4.7.9.1 Protocol Workload

For site 602591, the optimal protocol workloads were:

- To maximize clinical trial accrual
 - $2.5 < 3.0$ and $6.0 < 6.5$ protocols per staff member by study characteristic analysis (**Appendix F-7f**)
 - $2.5 < 3.0$ and $7.0 < 7.5$ protocols per FTE by study characteristic analysis (**Appendix F-7g**)
 - $6.5 < 7.0$ protocols per staff member by regression modeling (**Appendix F-8c**)
 - $7.0 < 7.5$ protocols per FTE by regression modeling (**Appendix F-8d**)
- To minimize the percentage of zero-accruing protocols

- 6.0 < 6.5 protocols per staff member by study characteristic analysis (**Appendix F-7f**)
- < 1 protocol per FTE by study characteristic analysis (**Appendix F-7g**)
- 6.5 < 7.0 protocols per staff member by regression modeling (**Appendix F-9c**)
- 7.0 < 7.5 protocols per FTE by regression modeling (**Appendix F-9d**)
- To maximize the percentage of protocols accruing at least four subjects
 - 6.0 < 6.5 protocols per staff member by study characteristic analysis (**Appendix F-7f**)
 - 2.5 < 3.0 protocols per FTE by study characteristic analysis (**Appendix F-7g**)
 - No significant association per staff member/FTE by regression modeling (**Table 4-32**)
- To minimize time to activate a protocol
 - < 1 protocol per staff member/FTE by study characteristic analysis (**Appendix F-7f** and **Appendix F-7g**)
 - < 1 protocol per staff member/FTE by regression modeling (**Appendix F-11b** and **Appendix F-11c**)

4.7.9.2 Accrual Prediction

Disease teams predicted a total of 6,299 subjects, 298% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the disease team correctly predicted it would 100% of the time. Disease teams correctly predicted 2.7% of the time that a study would not accrue four subjects. Total accuracy of the

disease team on this binary measure was 45.8% (**Table 4-52**; page 254). Disease teams predicted the category of actual accrual 17.0% of the time (**Appendix G-3i**).

The site-specific prediction model predicted a total of 2,177 subjects, 103% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the site-specific prediction model correctly predicted that studies would accrue at least four subjects 70.9% of the time. The model correctly predicted 77.3% of the time that a study would not accrue four subjects. Total accuracy of the site-specific model was 74.4% (**Table 4-52**; page 254). The site-specific model predicted the category of actual accrual 41.0% of the time (**Appendix G-2i**). When the trial actually accrued at least four subjects, both the disease team and the site-specific model predicted it would 70.9% of the time.

The unadjusted, overall prediction model predicted a total of 2,644 subjects, 125% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 82.9% of the time. The model correctly predicted 55.5% of the time that a study would not accrue four subjects. Total accuracy of the unadjusted, overall model was 67.6% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 35.2% of the time (**Appendix G-4i**).

The adjusted, overall prediction model predicted a total of 2,762 subjects, 131% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would

be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 85.1% of the time. The model correctly predicted 50.0% of the time that a study would not accrue four subjects. Total accuracy of the adjusted, overall model was 65.6% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 33.2% of the time (**Appendix G-5i**).

4.7.10: Site 689326

The number of protocols submitted was 456, contributing 7.9% to the total number of protocols studied in this project. Of these, no studies (0.0%) accrued zero subjects. The number of protocols that accrued at least four subjects was 324 (71.1%). Total contributed accrual was 6,705 subjects (13.6% of overall accrual). The distribution of protocol accrual is displayed in **Appendix E-1j**.

Activation times, usage of a feasibility committee, and regulatory functions of staff within and/or external to the centralized clinical trials office were not provided by the site.

4.7.10.1 Accrual Prediction

Disease teams predicted a total of 18,834 subjects, 281% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the disease team correctly predicted it would 100% of the time. Disease teams correctly predicted 1.5% of the time that a study would not accrue four subjects. Total accuracy of the

disease team on this binary measure was 71.5% (**Table 4-52**; page 254). Disease teams predicted the category of actual accrual 19.7% of the time (**Appendix G-3j**).

The site-specific prediction model predicted a total of 6,470 subjects, 96% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the site-specific prediction model correctly predicted that studies would accrue at least four subjects 95.4% of the time. The model correctly predicted 32.6% of the time that a study would not accrue four subjects. Total accuracy of the site-specific model was 77.2% (**Table 4-52**; page 254). The site-specific model predicted the category of actual accrual 35.5% of the time (**Appendix G-2j**). When the trial actually accrued at least four subjects, both the disease team and the site-specific model predicted it would 95.4% of the time.

The unadjusted, overall prediction model predicted a total of 4,421 subjects, 66% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 83.3% of the time. The model correctly predicted 53.8% of the time that a study would not accrue four subjects. Total accuracy of the unadjusted, overall model was 74.8% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 37.1% of the time (**Appendix G-4j**).

The adjusted, overall prediction model predicted a total of 4,582 subjects, 68% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would

be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 86.7% of the time. The model correctly predicted 48.5% of the time that a study would not accrue four subjects. Total accuracy of the adjusted, overall model was 75.7% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 36.2% of the time (**Appendix G-5j**).

4.7.11: Site 696337

The number of protocols submitted was 182, contributing 3.1% to the total number of protocols studied in this project. Of these, 38 (20.9%) accrued zero subjects. The number of protocols that accrued at least four subjects was 77 (42.3%). Total contributed accrual was 831 subjects (1.7% of overall accrual). The distribution of protocol accrual is displayed in **Appendix E-1k**.

The average time to activate a study was 155.6 ± 114.2 days (median 137 days). This was faster than the overall activation time of 190.1 ± 164.0 days (median 154 days). This site instated a feasibility committee in July 2013 for all studies. The site reported a centralized clinical trials office with regulatory functions being consolidated to the centralized office throughout the study period. Staff within the centralized clinical trials office are dedicated solely to regulatory functions, therefore staff workload did not differ between protocols per staff member and protocols per FTE. The distribution of protocol workload by staffing model and outcomes are diagramed in **Appendix F-1h** through **F-6h**.

4.7.11.1 Protocol Workload

For site 696337, the optimal protocol workloads were:

- To maximize clinical trial accrual
 - < 1 protocol and $8.0 < 8.5$ protocols per staff member/FTE by study characteristic analysis (**Appendix F-7h**)
 - No significant association per staff member/FTE by regression modeling (**Table 4-32**)
- To minimize the percentage of zero-accruing protocols
 - < 1 protocol per staff member/FTE by study characteristic analysis (**Appendix F-7h**)
 - No significant association per staff member/FTE by regression modeling (**Table 4-32**)
- To maximize the percentage of protocols accruing at least four subjects
 - < 1 protocol per staff member/FTE by study characteristic analysis (**Appendix F-7h**)
 - No significant association per staff member/FTE by regression modeling (**Table 4-32**)
- To minimize time to activate a protocol
 - $8.0 < 8.5$ protocols per staff member/FTE by study characteristic analysis (**Appendix F-7h**)
 - $1.5 < 2.0$ protocols per staff member/FTE by regression modeling (**Appendix F-11d**)

4.7.11.2 Accrual Prediction

Disease teams predicted a total of 1,917 subjects, 231% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the disease team correctly predicted it would 100% of the time. Disease teams correctly predicted 7.6% of the time that a study would not accrue four subjects. Total accuracy of the disease team on this binary measure was 46.7% (**Table 4-52**; page 254). Disease teams predicted the category of actual accrual 8.2% of the time (**Appendix G-3k**).

The site-specific prediction model predicted a total of 864 subjects, 104% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the site-specific prediction model correctly predicted that studies would accrue at least four subjects 72.7% of the time. The model correctly predicted 81.9% of the time that a study would not accrue four subjects. Total accuracy of the site-specific model was 78.0% (**Table 4-52**; page 254). The site-specific model predicted the category of actual accrual 40.7% of the time (**Appendix G-2k**). When the trial actually accrued at least four subjects, both the disease team and the site-specific model predicted it would 72.7% of the time.

The unadjusted, overall prediction model predicted a total of 960 subjects, 116% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 71.4% of the time. The model correctly predicted 46.7% of the time that a study would not accrue four subjects. Total accuracy of the unadjusted, overall model was

57.1% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 24.7% of the time (**Appendix G-4k**).

The adjusted, overall prediction model predicted a total of 931 subjects, 112% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 67.5% of the time. The model correctly predicted 48.6% of the time that a study would not accrue four subjects. Total accuracy of the adjusted, overall model was 56.6% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 22.0% of the time (**Appendix G-5k**).

4.7.12: Site 714415

The number of protocols submitted was 529, contributing 9.1% to the total number of protocols studied in this project. Of these, 80 (15.1%) accrued zero subjects. The number of protocols that accrued at least four subjects was 297 (56.1%). Total contributed accrual was 5,530 subjects (11.2% of overall accrual). The distribution of protocol accrual is displayed in **Appendix E-11**.

The average time to activate a study was 184.4 ± 151.3 days (median 146 days). This was faster than the overall activation time of 190.1 ± 164.0 days (median 154 days). This site did utilize a feasibility committee. The site reported a centralized clinical trials office with some regulatory functions conducted through the office. Staff are dedicated solely to regulatory functions, therefore staff workload did not differ between protocols per staff member and protocols per

FTE. The distribution of protocol workload by staffing model and outcomes are diagrammed in **Appendix F-1g** through **F-6g**.

4.7.12.1 Protocol Workload

For site 714145, the optimal protocol workloads were:

- To maximize clinical trial accrual
 - 5.0 < 5.5 protocols per staff member/FTE by study characteristic analysis
(**Appendix F-7i**)
 - 13 < 14 protocols per staff member/FTE by regression modeling (**Appendix F-8e**)
- To minimize the percentage of zero-accruing protocols
 - 8.5 < 9.0 protocols per staff member/FTE by study characteristic analysis
(**Appendix F-7i**)
 - No significant association per staff member/FTE by regression modeling (**Table 4-32**)
- To maximize the percentage of protocols accruing at least four subjects
 - 5.0 < 5.5 protocols per staff member/FTE by study characteristic analysis
(**Appendix F-7i**)
 - No significant association per staff member/FTE by regression modeling (**Table 4-32**)
- To minimize time to activate a protocol

- < 1 protocol (only two protocols reported) or 10 < 11 (more than two protocols) per staff member/FTE by study characteristic analysis (**Appendix F-7i**)
- No significant association per staff member/FTE by regression modeling (**Table 4-32**)

4.7.12.2 Accrual Prediction

Disease teams predicted a total of 8,629 subjects, 156% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the disease team correctly predicted it would 98.0% of the time. Disease teams correctly predicted 19.0% of the time that a study would not accrue four subjects. Total accuracy of the disease team on this binary measure was 63.3% (**Table 4-52**; page 254). Disease teams predicted the category of actual accrual 29.3% of the time (**Appendix G-31**).

The site-specific prediction model predicted a total of 11,449 subjects, 207% of the number of subjects actually accrued to studies at the site. There was one outlier where the prediction model predicted accrual of over 3,000 subjects. Removing this outlier reduced the predicted total to 5,065 subjects (92% of actual accrual). In predicting whether at least four subjects would be accrued to a study, the site-specific prediction model correctly predicted that studies would accrue at least four subjects 81.5% of the time. The model correctly predicted 59.9% of the time that a study would not accrue four subjects. Total accuracy of the site-specific model was 72.0% (**Table 4-52**; page 254). The site-specific model predicted the category of actual accrual 40.3%

of the time (**Appendix G-21**). When the trial actually accrued at least four subjects, both the disease team and the site-specific model predicted it would 80.5% of the time.

The unadjusted, overall prediction model predicted a total of 8,697 subjects, 157% of the number of subjects actually accrued to studies at the site. There was one outlier where the prediction model predicted accrual of over 3,000 subjects. Removing this outlier reduced the predicted total to 4,745 subjects (86% of actual accrual). In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 83.5% of the time. The model correctly predicted 54.3% of the time that a study would not accrue four subjects. Total accuracy of the unadjusted, overall model was 70.7% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 35.0% of the time (**Appendix G-41**).

The adjusted, overall prediction model predicted a total of 8,943 subjects, 162% of the number of subjects actually accrued to studies at the site. There was one outlier where the prediction model predicted accrual of over 3,000 subjects. Removing this outlier reduced the predicted total to 4,907 subjects (89% of actual accrual). In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 85.2% of the time. The model correctly predicted 50.4% of the time that a study would not accrue four subjects. Total accuracy of the adjusted, overall model was 69.9% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 34.8% of the time (**Appendix G-51**).

4.7.13: Site 715532

The number of protocols submitted was 225, contributing 3.4% to the total number of protocols studied in this project. Of these, 66 (29.3%) accrued zero subjects. The number of protocols that accrued at least four subjects was 82 (36.4%). Total contributed accrual was 1,313 (2.7% of overall accrual). The distribution of protocol accrual is displayed in **Appendix E-1m**.

Activation times, usage of a feasibility committee, and regulatory functions of staff within and/or external to the centralized clinical trials office were not provided by the site.

4.7.13.1 Accrual Prediction

Disease teams predicted a total of 4,756 subjects, 362% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the disease team correctly predicted it would 98.8% of the time. Disease teams correctly predicted 18.1% of the time that a study would not accrue four subjects. Total accuracy of the disease team on this binary measure was 47.6% (**Table 4-52**; page 254). Disease teams predicted the category of actual accrual 10.2% of the time (**Appendix G-3m**).

The site-specific prediction model predicted a total of 1,409 subjects, 107% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the site-specific prediction model correctly predicted that studies would accrue at least four subjects 70.7% of the time. The model correctly predicted 81.8% of the time that a study would not accrue four subjects. Total accuracy of the site-specific model was 77.8%

(**Table 4-52**; page 254). The site-specific model predicted the category of actual accrual 48.4% of the time (**Appendix G-2m**). When the trial actually accrued at least four subjects, both the disease team and the site-specific model predicted it would 69.5% of the time.

The unadjusted, overall prediction model predicted a total of 1,903 subjects, 145% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 86.6% of the time. The model correctly predicted 43.4% of the time that a study would not accrue four subjects. Total accuracy of the unadjusted, overall model was 59.1% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 26.2% of the time (**Appendix G-4m**).

The adjusted, overall prediction model predicted a total of 1,218 subjects, 93% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 69.5% of the time. The model correctly predicted 72.0% of the time that a study would not accrue four subjects. Total accuracy of the adjusted, overall model was 71.1% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 34.7% of the time (**Appendix G-5m**).

4.7.14: Site 846594

The number of protocols submitted was 294, contributing 5.1% to the total number of protocols studied in this project. Of these, 64 (21.8%) accrued zero subjects. The number of protocols that accrued at least four subjects was 115 (39.1%). Total contributed accrual was 1,854 subjects (3.8% of overall accrual). The distribution of protocol accrual is displayed in **Appendix E-1n**.

Activation times, usage of a feasibility committee, and regulatory functions of staff within and/or external to the centralized clinical trials office were not provided by the site.

4.7.14.1 Accrual Prediction

Disease teams predicted a total of 6,586 subjects, 355% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the disease team correctly predicted it would 100% of the time. Disease teams correctly predicted 10.1% of the time that a study would not accrue four subjects. Total accuracy of the disease team on this binary measure was 45.2% (**Table 4-52**; page 254). Disease teams predicted the category of actual accrual 8.2% of the time (**Appendix G-3n**).

The site-specific prediction model predicted a total of 2,223 subjects, 120% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the site-specific prediction model correctly predicted that studies would accrue at least four subjects 76.5% of the time. The model correctly predicted 74.3% of the time that a study would not accrue four subjects. Total accuracy of the site-specific model was 75.2%

(**Table 4-52**; page 254). The site-specific model predicted the category of actual accrual 42.9% of the time (**Appendix G-2n**). When the trial actually accrued at least four subjects, both the disease team and the site-specific model predicted it would 76.5% of the time.

The unadjusted, overall prediction model predicted a total of 2,480 subjects, 134% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 85.2% of the time. The model correctly predicted 53.6% of the time that a study would not accrue four subjects. Total accuracy of the unadjusted, overall model was 66.0% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 32.0% of the time (**Appendix G-4n**).

The adjusted, overall prediction model predicted a total of 2,562 subjects, 138% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 85.2% of the time. The model correctly predicted 50.3% of the time that a study would not accrue four subjects. Total accuracy of the adjusted, overall model was 63.9% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 30.3% of the time (**Appendix G-5n**).

4.7.15: Site 997056

The number of protocols submitted was 93, contributing 1.6% to the total number of protocols studied in this project. Of these, 12 (12.9%) accrued zero subjects. The number of protocols that accrued at least four subjects was 50 (53.8%). Total contributed accrual was 754 subjects (1.5% of overall accrual). The distribution of protocol accrual is displayed in **Appendix E-1o**.

Activation times, usage of a feasibility committee, and regulatory functions of staff within and/or external to the centralized clinical trials office were not provided by the site.

4.7.15.1 Accrual Prediction

Disease teams predicted a total of 1,490 subjects, 198% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the disease team correctly predicted it would 100% of the time. Disease teams correctly predicted 4.7% of the time that a study would not accrue four subjects. Total accuracy of the disease team on this binary measure was 55.9% (**Table 4-52**; page 254). Disease teams predicted the category of actual accrual 20.4% of the time (**Appendix G-3o**).

The site-specific prediction model predicted a total of 752 subjects, 100% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the site-specific prediction model correctly predicted that studies would accrue at least four subjects 86.0% of the time. The model correctly predicted 55.8% of the time that a study would not accrue four subjects. Total accuracy of the site-specific model was 72.0%

(**Table 4-52**; page 254). The site-specific model predicted the category of actual accrual 38.7% of the time (**Appendix G-2o**). When the trial actually accrued at least four subjects, both the disease team and the site-specific model predicted it would 86.0% of the time.

The unadjusted, overall prediction model predicted a total of 697 subjects, 92% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 84.0% of the time. The model correctly predicted 48.8% of the time that a study would not accrue four subjects. Total accuracy of the unadjusted, overall model was 67.7% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 36.6% of the time (**Appendix G-4o**).

The adjusted, overall prediction model predicted a total of 724 subjects, 96% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 86.0% of the time. The model correctly predicted 46.5% of the time that a study would not accrue four subjects. Total accuracy of the adjusted, overall model was 67.7% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 40.9% of the time (**Appendix G-5o**).

4.7.16: Site 998666

The number of protocols submitted was 193, contributing 3.3% to the total number of protocols studied in this project. Of these, 55 (28.5%) accrued zero subjects. The number of protocols that accrued at least four subjects was 80 (41.5%). Total contributed accrual was 1,092 subjects (2.2% of overall accrual). The distribution of protocol accrual is displayed in **Appendix E-1p**.

This site did not utilize a feasibility committee. Activation times and regulatory functions of staff within and/or external to the centralized clinical trials office were not provided by the site.

4.7.16.1 Accrual Prediction

Disease teams predicted a total of 2,110 subjects, 193% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the disease team correctly predicted it would 95.0% of the time. Disease teams correctly predicted 19.5% of the time that a study would not accrue four subjects. Total accuracy of the disease team on this binary measure was 50.8% (**Table 4-52**; page 254). Disease teams predicted the category of actual accrual 19.7% of the time (**Appendix G-3p**).

The site-specific prediction model predicted a total of 1,154 subjects, 106% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the site-specific prediction model correctly predicted that studies would accrue at least four subjects 66.3% of the time. The model correctly predicted 77.0% of the time that a study would not accrue four subjects. Total accuracy of the site-specific model was 72.5%

(**Table 4-52**; page 254). The site-specific model predicted the category of actual accrual 37.3% of the time (**Appendix G-2p**). When the trial actually accrued at least four subjects, both the disease team and the site-specific model predicted it would 62.5% of the time.

The unadjusted, overall prediction model predicted a total of 1,219 subjects, 112% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 76.3% of the time. The model correctly predicted 50.4% of the time that a study would not accrue four subjects. Total accuracy of the unadjusted, overall model was 61.1% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 30.6% of the time (**Appendix G-4p**).

The adjusted, overall prediction model predicted a total of 1,054 subjects, 97% of the number of subjects actually accrued to studies at the site. In predicting whether at least four subjects would be accrued to a study, the unadjusted model correctly predicted that studies would accrue at least four subjects 67.5% of the time. The model correctly predicted 61.9% of the time that a study would not accrue four subjects. Total accuracy of the adjusted, overall model was 64.2% (**Table 4-52**; page 254). The unadjusted, overall model predicted the category of actual accrual 27.5% of the time (**Appendix G-5p**).

CHAPTER 5. DISCUSSION

5.1 Major Findings Related to Clinical Research Site Efficiency

The purpose of this dissertation project is to address the gap in knowledge regarding how to objectively describe and quantitate optimal study activation workload at the study site level while maximizing the probability of clinical trial success, defined as site clinical trial accrual.

To address these issues, this project examined three areas of clinical trial activation: the use of a feasibility committee for study selection (Specific Aim 1), regulatory workflow (Specific Aim 2), and study selection (Specific Aim 3).

This study showed that the addition of a feasibility committee, which is dedicated to determining whether the resources and ability to complete a clinical trial, statistically increased clinical trial accrual by just under one subject per protocol and decreased activation time. Regarding optimal protocol workload levels, no single definitive protocol workload was found that both minimized activation time and increased clinical trial accrual; however, regression analyses showed that a protocol workload of $3.5 < 5.0$ protocols per staff member maximized clinical trial accrual and workloads between 1.0 and 1.9 protocols minimized activation time.

In addition, this work presents a systematic method for developing a model that predicts clinical trial accrual that was more accurate than the disease teams at all 16 participating centers. Several characteristics were consistently significantly associated with clinical trial accrual. These included the protocol's sponsor type (externally peer reviewed, institutional, industry, or national group), the number of national sites participating in the study, the number of months a study was

accruing nationally, the number of months the study had already been accruing nationally prior to activating the study, and the disease team conducting the study. Overall, the accuracy of the site-specific accrual prediction models was 74.0% in determining whether a study would accrue at least four subjects and 38.1% accurate at the accrual category levels. In comparison, the current standard—the disease team or investigators conducting the study—were 57.3% accurate in determining whether a study would accrue at least four subjects and only 21.7% accurate at the accrual category level.

5.2 Context of the Problem and the Project

Numerous reports over the last 15 years have highlighted the increasing cost and time it takes to develop a drug and bring it to market. Of the approximate 11 years it takes to develop a drug, approximately eight years are spent in clinical testing. Despite the time dedicated to the clinical development of an agent, many clinical trials are not able to assess their endpoints due to being underpowered from low accrual. This lack of efficiency leads to research waste measurable in monetary, scientific, and personnel resources. This issue remains even though the majority of phase III clinical trials extend their enrollment periods and add sites beyond the number originally planned. At the site level, multiple reports indicate the prolonged process of activating a clinical trial and indicate a large number of studies not accruing at the site level.

Reform is needed to transform this paradigm in the clinical research enterprise; however, national change is slow. This requires institutions conducting clinical research to use effective

measures for conducting innovative clinical research within the current paradigm, which are not currently available and is the focus of this dissertation project.

5.3 Implications of Findings

5.3.1 Specific Aim 1

This work adds to the existing literature in that it provides preliminary evidence supporting the effectiveness of a feasibility committee at cancer centers to screen protocols with the ultimate goal of increasing clinical trial accrual. The use of a feasibility committee was found to increase clinical trial accrual. However, while studies averaged higher accrual when a feasibility committee was utilized, it should be noted that this study did not see a difference in the likelihood that a study would accrue zero subjects with the use of a feasibility committee. This suggests other factors besides a feasibility committee influence accrual. Clinical trial accrual may indirectly benefit from feasibility committee review as the purpose of the feasibility committee is to abandon protocols that will not succeed. Thus while the final portfolio may still have clinical trials that fail to enroll as planned, the volume of such trials should be lessened.

The impactful finding of this analysis regarding feasibility committee usage was the reduction in overall activation time, or the time it takes to get a protocol open to accrual at a site. This is counterintuitive as an additional process (feasibility committee review) is being added to the activation pipeline. However, thought out, it makes sense. The purpose of the feasibility committee, according to stage-gate framework, is to eliminate studies from the activation pipeline that will ultimately not succeed. This reduces the number of projects in the activation

pipeline. With fewer projects in the activation pipeline, more focus can be given to the projects within the pipeline and, therefore, they can be activated more quickly.

In the move to make research more participant-centric, adding patient advocates and members of the community to feasibility committees (as is often done with the IRBs) may assist in the accrual process. These advocates can inform the committee as well as the research center of research priorities of interest to the patient population and/or community, thus providing information regarding the potential population to draw subjects from. These advocates could inform the community of research going forward and provide another resource for recruiting subjects. Having community buy-in may also address an aspect of feasibility that may not be apparent within the institution, answering the question of “do people want to participate in this type of research study,” “people will be more/less likely to participate because of these features,” and other valuable insights to enhance study performance.

Despite existing requirements surrounding the conduct of clinical trials, there is no requirement or recommendation to utilize a feasibility committee, let alone any guidelines regarding the constitution, review criteria, or decision-making processes for feasibility committees. This project did not evaluate the composition of the feasibility committees at each institution nor did it examine the criteria that committees used to determine whether a project was feasible (or not feasible). Given the significant, yet modest, improvements noted in this study, future work is warranted to identify what components/characteristics of feasibility committee review maximize the positive outcomes (such as minimizing activation time or maximizing accrual). The results

of such work should inform the entities who grant monies to fund clinical trial infrastructure (e.g. the NCI Cancer Center Support Grant [CCSG] and NIH Clinical and Translational Science Award [CTSA]) who should, in turn, revise requirements to reflect best clinical research operations practice.

5.3.2 Specific Aim 2

This work also presents the first known analysis to quantitate optimal regulatory workload to minimize activation time and/or maximize clinical trial accrual. Previous work has explored the concept of protocol acuity, or determining how difficult a protocol is, which is hypothesized to correlate to how much work it takes to execute the protocol (139). However, work associated with this methodology was limited to the clinical coordination associated with clinical trials (e.g. subject interactions and visits) and did not address regulatory activities, including study activation.

As research requirements continue to become more rigorous, it has become more common to see entire job positions dedicated to the management of the regulatory requirements of a clinical trial, the regulatory coordinator. Regulatory coordinators are persons trained in the regulatory requirements surrounding clinical trial operations. These people are trained (and can be certified) in federal regulations regarding human subjects protection, conflict of interest regulations, Health Information Portability and Accountability Act (HIPAA) regulations, and Food and Drug Administration (FDA) requirements. While it is more common to see dedicated regulatory coordinators in larger institutions, such as those with a Cancer Center Support Grant

(CCSG), it is not universal. Another common model is a more comprehensive study approach, where the same person manages both the regulatory management of a study as well as the clinical coordination (e.g. consenting subjects, data management, coordinating monitor visits, etc.). Because of these two models, two definitions of workload were assessed. The first is number of full-time equivalents (FTE). This quantifies only the portion of the job description dedicated to regulatory work, allowing assessment of workload to be based solely on regulatory work and not other job duties (such as clinical coordination). The second method is to assess workload per staff member, regardless of the other job duties that may be involved.

This study did find a significant relationship between protocol workload (both protocols per staff member and FTE) and clinical trial accrual. While regulatory personnel do not directly accrue subjects, it is rational to consider how well a study accrues based on activation timeline as the longer it takes to open a study, the fewer available spots there are to fill (due to ongoing national accrual). Additionally, the science and novelty of a study ages over time, making it less attractive to investigators and/or prospective subjects. A statistically significant relationship was also found with activation time. However, there was not a single protocol workload that best minimized activation time while maximizing trial accrual. Optimal protocol workload for a single outcome (e.g. clinical trial accrual) was not consistent between sites. Patterns in the data emerged indicating that protocol workloads between 3.5 and 5.0 protocols per staff member/FTE consistently performed well on accrual outcomes while smaller workloads, between 1.0 and 2.0 protocols per staff member/FTE, optimized activation outcomes. While a protocol workload of $1.0 < 1.9$ protocols per staff member/FTE may not be realistic for a dedicated staffing model, it

may be more realistic for a mixed staffing model where the person who manages the regulatory aspects of the trial is also the clinical research coordinator, handling subject consent, visits, and data management in addition to regulatory functions. Based on the findings of this project, further exploration of this topic is warranted utilizing a study design that includes the whole activation workload.

5.3.2.1 Considerations for Future Work

This project did not see an inverse curvilinear relationship between regulatory workload and optimum productivity, as described by the MTM framework, nor was there a linear relationship. In completing this aim, several factors emerged that should be considered in future work on this issue. Most institutions reported dedicated regulatory staff, with only two of the seven contributing institutions reporting differences between the number of staff members and number of FTEs, therefore there were not many differences between the two types of protocol workload examined, as shown by the near-perfect Pearson correlation coefficient between these two measures (0.98). More careful calculation of the time spent on regulatory work as well as how many protocols were assigned per staff member should both be considered before concluding that the MTM framework does not fit the workload. In this project, the % FTE and number of staff members were reported by each center, but in the general classification of regulatory work, not subdivided down to the activation phase. In addition, general data on the number of people/FTE working on regulatory projects for the overall center were collected; actual workloads per person were not obtained. To better measure the % FTE of regulatory work, a method such as effort tracking should be used to quantitate the actual work associated with study

activation, and not other regulatory actions such as maintenance of ongoing research or IRB study closure. More precise data, such as the actual time devoted per project to study activation processes and the actual number of projects per staff member, would provide greater insight to what workload best optimizes activation time as well as whether the relationship between productivity and outcomes follow an inverse curvilinear relationship. Knowing that relationship and optimal workloads allows for clinical research administrators to allocate resources efficiently, reducing waste in both personnel and monetary resources.

5.3.4 Specific Aim 3

The third aspect of clinical trial activation examined was study selection. Specific Aim 3 examined the ability for past performance to predict future performance. Utilizing protocol and institutional factors known prior to study activation, we were able to utilize a systematic methodology to create regression models that predicted accrual better than the current standard, disease team prediction. Only one variable, number of national sites, was statistically significant at all 16 centers. Three variables (sponsor type, total number of months accruing nationally, and the number of months of national accrual already completed) were significantly related with site-specific accrual at 15 of 16 centers. Disease teams were significantly related to clinical trial accrual at 14 centers, but contrary to previous reports (102), not as strongly correlated with local accrual as the protocol variables listed above. Other protocol characteristics, such as randomized design, precision medicine status, and the inclusion of a placebo arm, while reported in the literature as affecting participation rates in clinical trials, were much more variable and site-specific on whether they were associated with site-level accrual. These protocol-level

characteristics were also variable on their impact on protocol workload, affecting sites and outcomes (accrual, activation times) inconsistently. These may represent factors that affect individual participant reasons for accrual (as previously reported), but at a global level do not impact the overall conduct of a clinical trial, just individual decision-making.

At all 16 centers, site-specific models utilizing associated protocol variables consistently outperformed the site disease teams in predicting clinical trial accrual, with no disease team able to more accurately predict accrual at either the number of accrual categories matched or binary value of the protocol able to accrue at least four subjects. Using variables consistently associated with site accrual, an overall model was developed to assess whether local site accrual could be accurately predicted utilizing a national model. While the overall model did not perform as well as the site-specific models, it was consistently better than disease team predictions. Disease team predictions had low specificity (under 30% at every institution). The low specificity measures how often the team would state that they would accrue under four subjects when they actually did accrue under four subjects. Partnered with the near 100% sensitivity rates, this highlights the optimism of the disease teams in their performance and ability to accrue subjects to trial.

This work presents a model that can be utilized within the feasibility review process to objectively determine clinical trial accrual prior to committing the significant resources to activate and conduct a trial. Barnard presented a framework for predictive models (47). The generated site-specific prediction models fit this framework, in that it is:

- 1) Simple to use and understand. The outcome of accrual is a meaningful metric in the clinical research enterprise. The site-specific model generates this value.
- 2) Able to adapt to epidemiologic and environmental changes. This project established a systematic methodology to determine which protocol and institutional variables to include. As these change, the model can be adapted to accommodate those changes.
- 3) Able to inform in commissioning decisions. Clinical trial administration and/or principal investigators can utilize this number in a feasibility review to determine whether undertaking the proposed study will yield desired results, whether they be purely accrual-related or budgetary.

In addition, the development of a systematic methodology to predict clinical trial accrual prior to study activation at the site level is a novel contribution to the literature. A search of the literature found only one paper that assessed the ability of a prospective model to accurately predict clinical trial accrual (106). This paper focused on participant characteristics and protocol characteristics (eligibility criteria) and assessed the availability of a patient population to meet clinical trial accrual. The majority of papers assessing clinical trial accrual at both the site and national levels looked at completed studies, one at a time, to determine what the characteristics of the enrolled subject population were and/or what the characteristics of people who refused to consent were. Other studies focused on the rate of accrual by looking at studies after they opened to accrual (past the “point of no return”) but while they were still accruing subjects. Statistical models were generated to determine whether a study would meet the accrual goal at the national level, and when that goal would be met. While these are very valuable contributions

and needed studies to address the issue of clinical trial accrual, it is reactive, looking at actual accrual patterns of ongoing or completed studies to determine how accrual can be improved within the same study or presenting the “what should have been done” scenario. For studies that are still ongoing, changes can be made to increase accrual, but the opportunities to increase accrual are fewer as the study has reached the “point of no return” by opening to accrual. Abilities to increase accrual are limited to protocol modifications, such as select eligibility criteria changes, that do not affect the integrity of the data collected to date. Additional participant outreach can be conducted to identify other locations of potential participants. However, gains in accrual and/or accrual rates are likely to be small. Significant resources have already been invested by the site to activate the study. Thus, difficult decisions to either invest more in a failing study or to abandon the study and waste the resources invested to date must be made.

This work examines a perspective of clinical trial execution not thoroughly explored, the site’s contribution to national trial efficiency. Current studies, including one published within the last six months (85), model clinical trial accrual at a national level to determine if a study, overall, will meet its national enrollment goal or measured enrollment rates after trial activation. As previously reported, up to 90% of trials extend enrollment periods and add additional sites to attempt to meet the enrollment goal. If sites had a mechanism to objectively choose which studies would be a success, then only studies that will accrue a needed number of subjects would be opened and sponsors could be given more accurate projections of the deliverable number of subject accruals. From the site perspective, knowledge of optimal workloads and prediction

modeling for clinical trials offers mechanisms to improve site efficiency. Staffing resources can be better planned and studies that will not meet a certain accrual requirement can be avoided. National clinical trial planning would come from the bottom-up, from the sites. This drives national change without needing to change national policy, but could be done simultaneously to more widespread regulatory and policy assessment.

5.3.5 Rare Disease and Pediatric Studies

The value used for the dichotomous cut-off was set at four for this project. However, this is a somewhat arbitrary number and can be adjusted based on institutional and program expectations. In the case of rare disease and pediatric studies (which, by definition, are rare diseases), a smaller number defining success may be justified. For these projects, it may be more appropriate to utilize a value of one or two, or perhaps utilize the predictive model to see how many subjects are likely to be accrued and then prioritize studies based on the predictive value. Another form of rare disease study is precision medicine. The inclusion (or exclusion) of subjects based on molecular subtypes often defines a very small eligible population, as mutation rates for many genes are under 5-10% prevalence and not uncommon to be less than 2%.

This study did not see a significant difference in accrual for precision medicine trials and non-precision medicine trials. Only 4.4% of the protocols in this study population were precision medicine protocols, or having a molecular target within the inclusion/exclusion criteria. The time period for which studies closed to accrual (2009 through the first part of 2015) represents a time where precision medicine was being designed and implemented; therefore, not many studies

meeting the definition of precision medicine had been completed. As studies initiate and complete under the president's Precision Medicine Initiative (PMI) and the vice president's "moonshot" mission, more information regarding these studies will become available for analysis.

Regarding feasibility for rare disease and pediatric studies, these studies are omitted under the Cancer Center Support Grant (CCSG) guidelines from undergoing low accrual monitoring once activated (as they are expected to have low accrual and accrue slowly). Thus it is prudent to determine how feasibility committees should proceed regarding these studies. A blanket exception from review may be warranted, but what may be a better approach is determining if the patient population even exists to open the study. If there is an available population, the accrual prediction model can be used to prioritize study activation. This would then bypass the "Go/Abandon" determination under the stage-gate framework and replace it with a "Go Now/Go Later/Abandon" gradient.

There may be differing factors affecting accrual for rare disease studies. When more precision medicine studies have been completed, the effect of accrual on factors such as specific genetic mutation or, more generally, prevalence of mutation rate may provide additional information to predict accrual to these types of protocols. Likewise, as pediatric cancer is a rare disease, different factors may be associated with accrual in this population of studies. Further research with a larger number of protocols is needed to better define these studies regarding accrual and workload.

5.3.6 Characteristics of Zero-Accruing Protocols

Within this study 1,053, or 18.2%, of clinical trials did not accrue any subjects at the local level. These trials were characteristically different than trials that accrued at least one subject (**Table 4-4**; page 138). In this sample population, trials that did not accrue any subjects were more likely to be randomized, a higher phase trial, have an efficacy endpoint, enroll pediatric subjects, have between 10 and 199 sites enrolling nationally, and be open to accrual longer at the national level. These factors align with what has been previously published as factors that affect clinical trial accrual, specifically phase (79-81), randomization (36, 80), and the number of participating sites (83). Factors that were not different in this study, but have been reported in other studies to differ, includes the inclusion of a placebo or observation arm (36, 82, 86). The national enrollment goal was also found to be significantly different (80), with accruing studies having a higher mean enrollment goal; however, the median study size was virtually the same (104 versus 108 for zero-accruing versus accruing studies, respectively). Thus this difference represents the difference in the distribution of sample sizes for studies that accrued versus those that did not, not necessarily that smaller studies do not accrue. Partnered with the number of national sites, this variable becomes more meaningful. Studies with larger numbers of sites were more likely to accrue zero subjects. When examining the two variables together and looking at the average number of anticipated accruals per site, studies that had zero accrual had a lower anticipated average number of subjects per site (8.9 versus 12.1; $p=0.02$). Given that zero-accruing studies had a higher proportion of studies with higher site numbers, this indicates that a higher number of sites were needed to meet the enrollment goal, anticipating low accrual per site. At the site

level, lower accrual was also anticipated, as disease teams, on average, predicted higher subject accrual for studies that accrued than those that did not accrue.

5.3.7 Additional Potential Associated Factors

This study was a secondary data analysis of existing data. As such, not all variables of interest were available for examination. In addition, it is quite likely that there are factors that affect accrual and activation time that were not included in the model. With the small sample size, it was not possible to examine whether NCI designation made a difference in either outcome; however, future work should explore this and other site characteristics, such as volume indicators and geographic variation. Potential variables to explore include the available patient population from which to draw subjects, the center's catchment area (area and/or population), or the number of competing centers within the catchment area. Indicators of research productivity, such as NIH and/or NCI funding or the number of subjects accrued in previous years, can be used to categorize sites as high/medium/low volume centers.

This study utilized the disease team, the management group of a clinical research protocol. These can be disease-driven (such as breast or lung cancer) or they can be area-focused (early phase protocols). The inclusion of this variable was intended to account for investigator and staff characteristics in a succinct manner. As these are institutionally determined, they do not consistently match across institutions. Therefore, this variable could not be included in the overall models. Exploring cross-cutting variables, such as disease/cancer, in addition to investigator characteristics, such as years of research experience, tenure status, or number of

protocols overseen, could provide more targeted estimates of trial accrual, especially in rare cancers such as sarcoma, which may be combined with other diseases, and as such may not be best estimated in these models.

This study focused on protocol and site characteristics and their association with clinical trial activation and accrual. However, as outlined in **Figure 1-1** (page 34), there are other categories of characteristics that can impact clinical trial efficiencies in both these areas, sponsor issues and participant issues. While sponsors cannot meet their accrual goals without sites, sites cannot meet their accrual goals without participants. Participant characteristics were not a focus of this project as these characteristics may not be predictable prior to study activation, though they nonetheless can impact the ability for a study to accrue. One set of characteristics that may greatly affect accrual, and has been studied in the literature, are inclusion/exclusion criteria. Participant characteristics may add additional information to the regression models and their ability to 1) predict ultimate accrual and 2) inform on the actual impact of feasibility committees and/or workload.

5.4 Limitations

As this project was based in the United States, the findings are likely limited to cancer centers based in the U.S. Additionally, the majority of participating cancer centers are associated with an academic institution (exceptions are the Mayo Clinic and Samuel Oschin Comprehensive Cancer Institute); therefore, the applicability of these results may not apply to free-standing, private-practice based cancer centers or hospital networks. Future work is needed to determine if

these centers have similar patterns regarding accrual. No cancer centers in the northeast participated in this project; therefore the results may not be generalizable to that geographic region of the United States. However, as the methodology for determining effective site-specific models consistently provided more accurate models at all 16 participating institutions and institutions in the northeast also adhere to the same NCI Cancer Center Support Grant (CCSG) guidelines, it is likely that this method will produce successful results at centers within this geographic region.

The protocols contributed by each institution only represent a portion of the total portfolio. Studies were limited to interventional clinical trials with a primary purpose of treatment or supportive care. Therefore, the prediction model does not cover projects with other primary purposes. Additionally, the total workload of each center is not represented. It is important to note that the complete study activation portfolio was not available for Specific Aim 2, as only studies that completed both the activation process and enrollment period were submitted for this project. Therefore a sensitivity analysis was performed. As the centralized clinical trial offices provided the data for this project, staffing information for units outside the office were not always available. When it was reported that only staffing numbers for the centralized clinical trial office were given, only those protocols covered by the office were included in the protocol workload analyses of Specific Aim 2. Thus, additional studies and more detailed staffing information are needed to assess optimal protocol workload.

Due to the skewed distribution of the dependent variable (clinical trial accrual) and the use of the negative binomial regression modeling, the pseudo- R^2 value cannot be interpreted as it can in linear regression. Despite this, it does act as a measure of completeness. The low pseudo- R^2 values and the less than 100% matching of actual data with predicted data indicate that the model may be missing variables.

Finally, the prediction model is not intended to assess scientific validity or value of a clinical trial. It should not be the only consideration in opening a clinical trial, but as one component in a larger feasibility analysis. In this study, all protocols had to undergo some level of scientific review, either at the national level (for externally peer reviewed projects and National Clinical Trial Network-sponsored protocols) or at the site (for institutional and industry projects), thus ensuring the integrity of the research conducted at cancer centers.

5.5 Policy Implications

This work provides preliminary evidence in support of feasibility committees to determine whether a study has the resources and ability to meet its goals. Future work should be conducted to determine what features and functions make these committees most effective, defined as maximizing predetermined outcomes (such as minimizing activation time and/or maximizing center/study accrual). Also examined in this project was the development of a systematic method to estimate clinical trial accrual prior to dedicating extensive resources to activating clinical trial workload. This methodology, in conjunction with existing participant identification and accrual prediction methods, such as those exemplified in London's work (106), can provide

objective data that can be considered by a feasibility committee when determining whether a clinical trial should be executed. As highlighted in the Lancet series, future research should focus on reducing research waste within clinical research operations. Future work should continue identifying factors and processes that contribute to successful clinical trial conduct (such as studies that meet accrual goals) as well as identify processes that detour clinical research from being conducted efficiently. Agencies who fund clinical research infrastructure, such as the NIH and NCI, should monitor research in this area and revise guidelines regarding the structure and use of feasibility committees as well as require objective decision-making processes for clinical trial selection in their funding requirements to ensure that institutions utilizing taxpayer dollars to conduct clinical research are doing so utilizing best practices.

5.6 Conclusions

Together, the information from each of the models resulting from the Specific Aims can be utilized to provide a multi-faceted analysis of the study activation and accrual pipelines of a site's clinical trial enterprise. These tools, utilized within a feasibility committee process, can inform an administrator and disease team whether the needed staffing levels to activate a protocol currently exist, as well as the probable accrual to a clinical trial. Harmonized with a site's scientific review process, this efficiency review can provide an objective process for improving efficiency at clinical research sites.

APPENDIX A – EMAIL RECRUITMENT LETTERS

First Email

Good evening,

At the University of Arizona Cancer Center, we have developed a regression model to prospectively predict how many people we can expect to accrue to a clinical trial (the outcome being the number of people accrued locally). The work was presented as a poster at the ASCO and AACI-CRI meetings this past summer (copies attached). We are now incorporating the model as part of our internal feasibility review at study start-up and actively validating it for our site. We have, so far, had very positive results that show that we can accurately foretell how many people we will accrue to a specific protocol for a majority (over 80%) of protocols.

Concurrently, I am a PhD candidate in Pharmaceutical Economics, Policy and Outcomes at the University of Arizona College of Pharmacy, and my dissertation topic is to expand this work to other cancer centers across the nation to assess how different (and alike) we are. I will be doing center specific models (which will be provided back to the centers that participate for their use) as well as an overall, hierarchical model to see if there are common data elements that can accurately predict accrual across sites.

Participating sites are asked to provide protocol level details regarding clinical trials that have completed enrollment over a 5-to-6-year period (January 2009 through current date, 2014). No patient-level data is requested. Data elements include local information, such as when the study opened and closed to accrual, how many people were accrued to trial locally, and what disease

team (or other designation of research management group) oversaw the protocol. Protocol level details include items such as: whether the study is a cooperative group protocol, phase of protocol, overall accrual goal, etc. An abstraction sheet has been developed by Ms. Tate and will be provided to participating sites.

Participating cancer centers will be coded with a six-digit code generated from a random number generator. This code will be used for identification purposes in all public presentations, reports (including Ms. Tate's dissertation) and publications regarding the study characteristic summaries, regression models and other statistical output. Each site will have knowledge of their code, but other participating sites (except for the University of Arizona) will not have access to the code list. A list of participating centers will be included in the dissertation paper and listed as a participating site in resulting peer-reviewed publications and conference presentations. Each center will be invited to co-author with the University of Arizona on resulting publications regarding the research. Members of the University of Arizona research team, specifically, the principal investigator/PhD student (Wendy Tate) and her dissertation committee will have access to the identifiable information for the purposes of data analysis and quality assurance.

If a center is interested in participating, I am happy to arrange further discussion to explore collaborative opportunities. I am very excited about this project and happy to talk to anyone at your center who may be interested in learning more about it.

Thank you,

Wendy

Attachments: 2014 AACI Abstract; 2014 AACI Poster

Second Email

Good morning,

It has been a couple of weeks since I sent the below email. I wanted to follow-up to see if there was any additional information I could provide regarding this project.

Thank you,

Wendy

Final Email

Good evening,

I wanted to follow up as it has been a few weeks since the last email I sent. I would love for your center to collaborate on this project, but I understand how busy things are. This will be my last email, so I want to take this opportunity to thank you for your time and consideration.

-Wendy

APPENDIX B – INTRODUCTION LETTER AND ABSTRACTION SHEET

Introduction Letter

Thank you for considering participating in this research project.

Preliminary data from our exploratory study shows that there are national and local study characteristics that predict clinical trial accrual at a single center. This proposed work seeks to expand this predictive model to ascertain whether this model can be applied nationally, both at the center level and overall at the national level. This work will be part of my dissertation project in Pharmaceutical Economics, Policy, and Outcomes. The single site report was presented as a poster at the ASCO 2014 Annual Meeting and the AACI-CRI meetings, both in Chicago, IL.

Working Title of Dissertation

Development of a national, multi-center, hierarchical linear regression model to predict overall and site-specific clinical trial protocol accrual at NCI designated cancer centers.

A Brief Description of the Project Background

In order for a clinical trial to draw scientifically sound conclusions from the data collected, the study must meet its accrual goal to be appropriately powered. According to an 2010 Institute of Medicine (IOM) report on the National Cancer Clinical Trial System for the 21st Century, 40% of cooperative group phase III trials opened between 2000 and 2007 did not meet their minimum accrual goal. Dilts et al. reported that 6.4% of Cancer Therapy Evaluation Program (CTEP), the

NCI program that vets NCI-sponsored clinical trials, did not accrue a single subject nationally (Dilts, 2010). Dilts et al. reported that, at four comprehensive cancer centers, approximately 20-40% of trials will not accrue any subjects locally (ibid). In addition, Dilts' report showed that it is more likely that an NCI-sponsored cooperative group trial (e.g. Southwest Oncology Group, or SWOG) will not accrue any subjects at the site than a study that is not sponsored by the cooperative group (38.8% versus 20.6%; ibid). The result is that many valuable scientific questions regarding the treatment of cancer go unanswered, as well an ongoing ethical debate of whether it was appropriate to place the subjects that did enroll in these trials at risk, as no benefit (the gain of scientific evidence for safety and/or effectiveness) was received.

A search of the literatures shows that several models have been created to assist in determining who will specifically join a clinical trial, whether a trial will meet its accrual goal based on its predicted accrual rate, and whether the trial as a whole will accrue the needed number of subjects. However, no papers model whether a single site will accrue a desirable number of subjects, defined as the number of subjects required to meet center-specific criteria regarding the economic needs of the center, scientific return to the center for conducting the trial locally, and ultimately, the center's contribution towards the NCI's mission to "assess the incorporation of state-of-the-art cancer treatments into clinical practice (NCI website)." The primary objective of the proposed study is to expand our current work, limited to a single center, expanding to a national level by applying a hierarchical linear model in order to predict clinical trial accrual at the overall protocol level, the level of the individual site, and the level of the site's disease team. The model will adjust for factors associated with the dependent variable (accrual), including the

NCI designation of the cancer center. With increasing cost to conduct clinical trials, it is imperative to select trials rationally for local activation. However, no models with the purpose of determining how well a site will accrue to trial exist.

If valid, this model would provide a quick, qualitative, objective and valuable metric to cancer center administration in assessing potential trial success as well as planning resource allocation and estimating costs. At the national level, entities such as the National Cancer Institute (NCI) can use this model when considering whether a trial can meet its recruitment endpoint as well as target sites that will best help meet the accrual goal.

Plan of Investigation and Methodologies to be Employed

All NCI designated cancer centers (41 comprehensive, 27 designated) will be approached for inclusion in this multi-center, retrospective cohort study. Permission will be sought from the Director and/or Associate Director of the Clinical Trials Unit. Study characteristic information regarding each participating Cancer Center will be obtained as well as at least five years of local (from the center's clinical trial management system) and national (from ClinicalTrials.gov) protocol-specific characteristics that will be used to define the protocols as well as model accrual.

Site Instructions

Complete the attached site study characteristic and abstraction sheets. The variables requested, as well as their definitions, are listed in the "Data Dictionary" tab. The requested time frame is January 1, 2009 through current day.

Data from each center will be modeled independently as well as be combined with all other centers using a negative binomial regression model, as the data will be count data and right skewed. Center models will utilize a multiple linear regression model while the overall model will utilize a hierarchical linear regression model, with the levels defined as the center, disease team, and finally, the protocol. Preliminary validation will be completed by modeling the cases that were used to create the model. A copy of each institution's model will be returned to the institution for evaluation and use as desired. Any publications regarding the model will mask the identity of the institution with recognition of the site's participation in the project only.

Thank you for your participation in this project. Please do not hesitate to contact me directly with any questions and/or comments.

Site Study characteristics tab

Site name:

Date completed:

Contact Person:

Role:

Phone:

Email:

NCI Designation: None/Cancer Center/Comprehensive

Setting: Academic/Community

Number of local sites reported in this data:

Data Dictionary Tab

Variable	Description	Type	Value	ClinicalTrials.gov field	Notes
ID	Protocol ID (unique identifier)	number			this does not have to be the protocol number, but can be another unique identifier used to distinguish one protocol from another
NCTID	ClinicalTrials.gov ID	text	NCTXXXXXXXX		this value is needed to check national protocol characteristics, if needed
Team	Disease team managing the protocol	text			this can be coded to mask the team's identities
TeamPredict	Disease team prediction of accrual	number			prior to study activation, prediction by the disease team of how many subjects will be accrued
Local_Accrual	Total local accrual obtained	number			
IND	Did the protocol have an IND	binary	0,1 (No/Yes)		
Phase	Phase of study	categorical	0,1,2,3,4,5,6,7,8,9 (Pilot, Phase 0, Phase I, Phase I/II, Phase II, Phase II/III, Phase III, Phase IV, Phase V, None)		
IRB	Institutional/local IRB used	binary	0,1 (No/Yes)		

OG	Cooperative group study	binary	0,1 (No/Yes)		any cooperative group (SWOG, GOG, COG, ECOG, CALGB, etc.)
Peds	Open to pediatric subjects	binary	0,1 (No/Yes)	Ages Eligible for Study	defined as less than 18 years of age
Random	Included a randomized component	binary	0,1 (No/Yes)	Purpose - Study Design	
Placebo	Included a placebo component	binary	0,1 (No/Yes)		specific statement of placebo as any part of an arm. This does NOT include standard of care, usual care or active treatment WITHOUT a placebo element
PriEnd	Primary endpoint	categorical	0,1,2 (Safety Only, Efficacy, Other)	Endpoint Classification	
Local_Open	Date opened to accrual locally	date	mm/dd/yyyy		
Local_Close	Date closed to accrual locally	date	mm/dd/yyyy		
NatEnroll	Expected national enrollment	number		Estimated Enrollment	from either the protocol or as stated on ClinicalTrials.gov
NatSites	Expected number of enrolling sites nationally	number		Contacts and Locations	from either the protocol or as stated on ClinicalTrials.gov
NatSitesCat	Expected number of enrolling sites nationally	categorical	0,1,2,3,4 (0-9 sites, 10-49 sites, 50-199 sites, 200+ sites, unknown)	Contacts and Locations	from either the protocol or as stated on ClinicalTrials.gov

Nat_Open	Date opened to accrual nationally	date	mm/dd/yyyy	Study Start Date	from local records or as stated on ClinicalTrials.gov (assuming a start on the first day of the month if from ClinicalTrials.gov)
Nat_Close	Date closed to accrual nationally (expected)	date	mm/dd/yyyy	Estimated Primary Completion Date	from local records or as stated on ClinicalTrials.gov (assuming closure on the first day of the month if from ClinicalTrials.gov)
Generated Variables					
Local_Days	Number of days opened to accrual locally	number	Local_Close - Local_Open		
DaysPer	Average number of days per accrual	number	Local_Days/Local_Accrual		
NonAccr	Study never accrued subjects	binary	0,1 (No/Yes)		
TotalMO	Total months opened to accrual nationally	number	(Nat_Close - Nat_Open)/30.4		
MOAccrDone	Months protocol opened nationally prior to local activation	number	(Local_Open - Nat_Open)/30.4		

APPENDIX C – CENTER INFRASTRUCTURE DATA COLLECTION SHEET

Infrastructure Recruitment Email

Thank you again for participating in this project. As the purpose of this project is to assess clinical trial study activation and maximize the potential to increase clinical trial accrual, I have expanded my specific aims to include a look at how feasibility committees and staffing levels impact study activation and, ultimately, correlate it to accrual. Attached, you will find another data collection sheet. I am kindly requesting this additional information so that study activation timelines, staff workload, and accrual can be explored to assess whether there is an optimal staffing model to minimize activation timelines and/or accrual as well as to assess whether a formal feasibility committee results in protocols that have higher per-protocol accrual.

I apologize for this additional request; however, in looking at the original data abstraction sheet, I see that the date that the protocol was started in the study activation process was not included (e.g. for OnCore, the date a protocol is listed as “New”). If possible, I would like to request this information for the protocols provided so that total study activation time can be calculated. As these data were collected several months ago, if you would like to include more up-to-date protocols, feel free to do so.

As always, if you have any questions or comments, please do not hesitate to contact me. Thank you again for your participation in this project.

Data Collection Sheet

Institution name:

Contact name:

Contact information:

Is an updated abstraction sheet being provided that includes the date that a protocol started the activation process (yes/no):

At what decision point is the protocol determined to be starting the study activation process (at what point is the "New" date inserted)

Please complete the table for each timepoint (January 2009, July 2009, January 2010, July 2010, January 2011, July 2011, January 2012, July 2012, January 2013, July 2013, January 2014, July 2014, January 2015, July 2015* (*only if updated protocol information is being submitted).

Institutional Infrastructure

At your center, is there a formal feasibility assessment performed for institutional protocols (separate from PRMS or IRB)

At your center, is there a formal feasibility assessment performed for industry protocols (separate from PRMS or IRB)

At your center, is there a formal feasibility assessment performed for national group protocols (separate from PRMS or IRB)

At your center, is there a formal feasibility assessment performed for externally peer reviewed protocols (separate from PRMS or IRB)

Is the formal feasibility assessment performed: prior to PRMS, concurrent to PRMS, after PRMS but before IRB, concurrent with IRB, after IRB, or N/A

Does your institution allow concurrent PRMS and IRB submission

Regulatory/Study Activation Staffing

Approximate percent of protocols with regulatory functions overseen by the centralized clinical trials office (yes/no)

Dedicated Study Activation Staff (yes/no/mixed)

Number of Study Activation Staff

Amount of Study Activation FTE

Dedicated Regulatory Staff (yes/no/mixed)

Number of Dedicated Regulatory Staff

Amount of Dedicated Regulatory FTE

Number of combined CRC/Regulatory Staff

Amount of combined CRC/Regulatory FTE

Number of combined Budget/Regulatory Staff

Amount of combined Budget/Regulatory FTE

For each disease management group, how many people (not FTE) are working on regulatory functions (Insert each disease team on a line - e.g. Breast, GI, Early Phase)

Data Dictionary

Term	Response	Definition	"Yes"	"No"	"Mixed"
Formal feasibility committee		a group of people whose responsibility with this committee (whether virtual or in-person) is to determine whether the staff or institutional resources are available to conduct a study			
Dedicated Study Activation Staff	Yes/No/Mixed	staff members who only work on study activation for clinical trials (e.g. no study maintenance)	all regulatory staff are dedicated solely to study activation functions	no regulatory staff are dedicated solely to study activation functions	some regulatory staff are dedicated solely to study activation functions and some have maintenance functions
Number of Dedicated Regulatory Staff	number	Number of people (regardless of FTE) that solely work on regulatory functions			
Amount of Dedicated Regulatory FTE	number	Amount of full-time effort (FTE; regardless of number of people) that is dedicated to regulatory functions			
Dedicated Regulatory Staff	Yes/No/Mixed	staff members who only work on regulatory functions for clinical trials (e.g. not a combined clinical research coordinator/regulatory job or budget/regulatory job)	all regulatory staff are dedicated solely to regulatory functions	no regulatory staff are dedicated solely to regulatory functions	some regulatory staff are dedicated solely to regulatory functions and some have other

					functions in addition to regulatory functions
Number of Dedicated Regulatory Staff	number	Number of people (regardless of FTE) that solely work on regulatory functions			
Amount of Dedicated Regulatory FTE	number	Amount of full-time effort (FTE; regardless of number of people) that is dedicated to regulatory functions			

APPENDIX D – FINAL DATABASE DATA DICTIONARY

Dependent variables

- LocalAccrual: Clinical trial accrual (from here forward, notated as “accrual”): total number of human subjects that will enroll on a clinical trial (format: continuous)
- ActivationTime: Total number of days between the start of activation and start of accrual for a protocol at a center (format: continuous)
 - Calculation: (Local_Open – Local_New)

Independent variables

- Phase: Study Phase (format: ordinal)
 - 0: Pilot
 - 1: Phase 0
 - 2: Phase I
 - 3: Phase I/II
 - 4: Phase II
 - 5: Phase II/III
 - 6: Phase III
 - 7: Phase IV
 - 9: None
 - 99: Missing
- DiseaseTeam: Primary management group/disease team responsible for study oversight.
The values of this variable differ between institutions (format: character)

- SponsorType: NCI category of protocol sponsor, as defined by the CCSG Data Table Guide (format: categorical; excerpted below)
 - 0: Externally Peer-Reviewed: R01s, SP0RES, U01s, U10s, P01s, CTEP, or any other clinical research study mechanism supported by the NIH or organizations on this list:

<http://cancercenters.cancer.gov/documents/PeerReviewFundingOrganizations508C.pdf>
 - 1: Institutional: In-house clinical research studies authored or co-authored by Cancer Center investigators and undergoing scientific peer review solely by the Protocol Review and Monitoring System of the Cancer Center. The Cancer Center investigator has primary responsibility for conceptualizing, designing, and implementing the clinical research study and reporting results.
 - 2: Industrial: A pharmaceutical company controls the design and implementation of these clinical research studies.
 - 3: National: NCI National Clinical Trials Network (NCTN) and other NIH-supported National Trial Networks
- Peds: Inclusion of pediatric population (format: binary; Y/N)
- Random: Use of randomized design (format: binary; Y/N)
- Placebo: Use of placebo (format: binary; Y/N)
- PriEnd: Primary endpoint of clinical trial (format: categorical)
 - 0: safety
 - 1: efficacy or safety&efficacy

- 2: other
- NatEnroll: National enrollment goal for protocol (format: continuous)
- NatSites: Number of participating sites globally (format: categorical)
 - 0: 1-9 sites
 - 1: 10-49 sites
 - 2: 50-199 sites
 - 3: 200+ sites
 - 4: unknown
- PrimaryPurpose: NCI category, restricted by inclusion criteria to therapeutic or supportive care, interventional protocols, as defined by the NCI Data Table Guide, version July 29, 2013; (135; format: binary)
 - 0: Supportive Care
 - 1: Treatment
- Precision: defined as a protocol having at least one inclusion or exclusion criterion in the ClinicalTrials.gov record that included at least one specific genetic aberration (mutation, etc.) determined by sequencing (not histology, immunohistochemistry, cytology, karyotyping, etc.) (format: categorical)
 - 0: Not precision medicine
 - 1: Precision medicine
 - 2: Conditional precision medicine
 - at least one of the following criterion were present: 1) if mutational status is known, it must be disclosed but does not disqualify a person from

participating; or 2) the mutational status does exclude if known, but testing is not required to participate; or 3) the mutational status does include/exclude but is one of a list of criteria that are annotated by an “or” and not all the criteria are mutation-based.

- IRB: IRB of record for the study (format: binary)
 - 0: External IRB
 - 1: Local IRB

Descriptor variables

- SiteID: random six-digit code to uniquely identify a site (format: categorical)
- ID: unique, center-specific protocol ID (format: character)
- NCTID: CT.gov number (format: string; 12 characters)
- Title: Short Title of study (format: character; max: 600 characters)
- Nat_Open: National study start date (format: date; MM/DD/YYYY)
- Nat_Close:
 - For studies closed to accrual nationally, abstracted actual closure date (format: date; MM/DD/YYYY)
 - For studies without a national closure date: Estimated national primary endpoint completion date (format: date; MM/DD/YYYY)
- Local_New: date that the study activation process was started for a protocol at a center (format: date; MM/DD/YYYY)

- Local_Open: earliest date that a protocol was open for accrual at a center (format: date; MM/DD/YYYY)
- Local_Close: date that a protocol was permanently closed to accrual at a center (format: date; MM/DD/YYYY)
- Designation: NCI designation of the center (format: categorical)
 - 0: Comprehensive cancer center
 - 1: Clinical cancer center
 - 2: Not NCI-designated
- Feasibility: usage of a feasibility committee, defining stage-gate process (format: binary)
 - 0: No feasibility committee
 - 1: Feasibility committee
- Reg_Staff: Protocol workload measure. Number of staff members working on regulatory functions (format: continuous)
- Reg_FTE: Protocol workload measure. Number of FTE working on regulatory functions (format: continuous)
- TeamPredict: Disease team prediction of accrual for a protocol at a center (format: continuous)
- SitePredict: Site-specific model prediction of accrual for a protocol at a center (format: continuous)
- Overall_Model: Overall, unadjusted model prediction of accrual for a protocol at a center (format: continuous)

- RE_Overall_Model: Overall, adjusted model prediction of accrual for a protocol at a center (format: continuous)

Calculated variables

- Local_Days: number of days a study was open to accrual at a center (format: continuous)
 - Local_Close – Local_Open
- DaysPer: average number of days per accrual at a center (format: continuous)
 - LocalAccrual/Local_Days
- TotalMO: Total number of months that accrual is expected to be ongoing nationally (format: continuous)
 - calculation: Nat_Close – Nat_Open
- MOAccrDone: Total number of months of accrual completed nationally prior to the center opening the study to accrual (format: continuous)
 - calculation: (Local_Open – Nat_Open)/30.4
 - 1 year = 365.25 days (accounting for ¼ day each year for leap year)
 - Calculate days per month as $365.25/12 = 30.4$
- Team_to_Actual: proportion of actual accrual estimated by the disease team (format: continuous)
 - calculation: TeamPredict/LocalAccrual
- SitePredict_to_Actual: proportion of actual accrual estimated by the site-specific model (format: continuous)
 - calculation: SitePredict/LocalAccrual

- OverallPredict_to_Actual: proportion of actual accrual estimated by the overall, unadjusted model (format: continuous)
 - calculation: Overall_Model/LocalAccrual
- RE_OverallPredict_to_Actual: proportion of actual accrual estimated by the overall, adjusted model (format: continuous)
 - calculation: RE_Overall_Model/LocalAccrual
- ActualAccrualCat: Actual accrual per protocol (format: categorical)
 - 0: < 1 subject
 - 1: 1 < 4 subjects
 - 2: 4 < 7 subjects
 - 3: 7 < 10 subjects
 - 4: 10 < 20 subjects
 - 5: 20 < 50 subjects
 - 6: 50+ subjects
- TeamPredictCat: Disease team predicted accrual per protocol (format: categorical)
 - 0: < 1 subject
 - 1: 1 < 4 subjects
 - 2: 4 < 7 subjects
 - 3: 7 < 10 subjects
 - 4: 10 < 20 subjects
 - 5: 20 < 50 subjects
 - 6: 50+ subjects

- SitePredictCat: Site-specific model predicted accrual per protocol (format: categorical)
 - 0: < 1 subject
 - 1: 1 < 4 subjects
 - 2: 4 < 7 subjects
 - 3: 7 < 10 subjects
 - 4: 10 < 20 subjects
 - 5: 20 < 50 subjects
 - 6: 50+ subjects

- OverallPredictCat: Overall, unadjusted model predicted accrual per protocol (format: categorical)
 - 0: < 1 subject
 - 1: 1 < 4 subjects
 - 2: 4 < 7 subjects
 - 3: 7 < 10 subjects
 - 4: 10 < 20 subjects
 - 5: 20 < 50 subjects
 - 6: 50+ subjects

- RE_OverallPredictCat: Overall, adjusted model predicted accrual per protocol (format: categorical)
 - 0: < 1 subject
 - 1: 1 < 4 subjects
 - 2: 4 < 7 subjects

- 3: $7 < 10$ subjects
 - 4: $10 < 20$ subjects
 - 5: $20 < 50$ subjects
 - 6: 50+ subjects
- Actual_Accrual_Binary: Indicates whether a protocol actually accrued at least four subjects (format: binary)
 - 0: < 4 subjects
 - 1: 4+ subjects
- TeamPredict_Binary Indicates whether the disease team predicted that a protocol would accrue at least four subjects (format: binary)
 - 0: < 4 subjects
 - 1: 4+ subjects
- SitePredict_Binary Indicates whether the site-specific model predicted that a protocol would accrue at least four subjects (format: binary)
 - 0: < 4 subjects
 - 1: 4+ subjects
- OverallPredict_Binary Indicates whether the overall, unadjusted model predicted that a protocol would accrue at least four subjects (format: binary)
 - 0: < 4 subjects
 - 1: 4+ subjects
- RE_OverallPredict_Binary Indicates whether the overall, adjusted model predicted that a protocol would accrue at least four subjects (format: binary)

- 0: < 4 subjects
 - 1: 4+ subjects
- Activation_Period: Six month period in which the Local_New date occurred (format: categorical)
 - 0: Jan-June 2007
 - 1: July-Dec 2007
 - 2: Jan-June 2008
 - 3: July-Dec 2008
 - 4: Jan-June 2009
 - 5: July-Dec 2009
 - 6: Jan-June 2010
 - 7: July-Dec 2010
 - 8: Jan-June 2011
 - 9: July-Dec 2011
 - 10: Jan-June 2012
 - 11: July-Dec 2012
 - 12: Jan-June 2013
 - 13: July-Dec 2013
 - 14: Jan-June 2014
 - 15: July-Dec 2014
 - 16: Jan-June 2015
 - 17: July-Dec 2015

- Prot_Active: Count of protocols with a Local_New date in the specified activation period
(format: continuous)
- Prot_Per_Staff: Average number of protocols in the activation period per staff member
(format: continuous)
 - calculation: $\text{Prot_Active}/\text{Reg_Staff}$
- Prot_Per_FTE: Average number of protocols in the activation period per FTE (format: continuous)
 - calculation: $\text{Prot_Active}/\text{Reg_FTE}$
- Prot_Per_Staff_Cat: Average number of protocols in the activation period per staff member (format: categorical)
 - 0: < 1 protocol
 - 1: 1.0 < 1.5 protocols
 - 2: 1.5 < 2.0 protocols
 - 3: 2.0 < 2.5 protocols
 - 4: 2.5 < 3.0 protocols
 - 5: 3.0 < 3.5 protocols
 - 6: 3.5 < 4.0 protocols
 - 7: 4.0 < 4.5 protocols
 - 8: 4.5 < 5.0 protocols
 - 9: 5.0 < 5.5 protocols
 - 10: 5.5 < 6.0 protocols
 - 11: 6.0 < 6.5 protocols

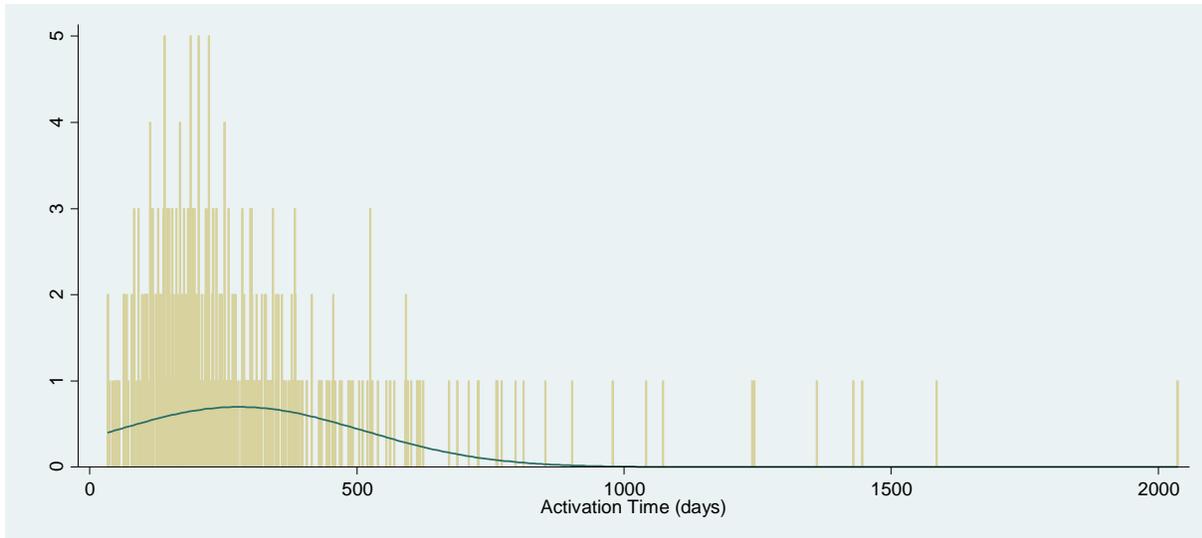
- 12: 6.5 < 7.0 protocols
 - 13: 7.0 < 7.5 protocols
 - 14: 7.5 < 8.0 protocols
 - 15: 8.0 < 8.5 protocols
 - 16: 8.5 < 9.0 protocols
 - 17: 9.0 < 9.5 protocols
 - 18: 9.5 < 10 protocols
 - 19: 10 < 11 protocols
 - 20: 11 < 12 protocols
 - 21: 12 < 13 protocols
 - 22: 13 < 14 protocols
 - 23: 14 < 15 protocols
 - 24: 15 < 16 protocols
 - 25: 16 < 17 protocols
- Prot_Per_FTE_Cat: Average number of protocols in the activation period per FTE
(format: categorical)
 - 0: < 1 protocol
 - 1: 1.0 < 1.5 protocols
 - 2: 1.5 < 2.0 protocols
 - 3: 2.0 < 2.5 protocols
 - 4: 2.5 < 3.0 protocols
 - 5: 3.0 < 3.5 protocols

- 6: 3.5 < 4.0 protocols
- 7: 4.0 < 4.5 protocols
- 8: 4.5 < 5.0 protocols
- 9: 5.0 < 5.5 protocols
- 10: 5.5 < 6.0 protocols
- 11: 6.0 < 6.5 protocols
- 12: 6.5 < 7.0 protocols
- 13: 7.0 < 7.5 protocols
- 14: 7.5 < 8.0 protocols
- 15: 8.0 < 8.5 protocols
- 16: 8.5 < 9.0 protocols
- 17: 9.0 < 9.5 protocols
- 18: 9.5 < 10 protocols
- 19: 10 < 11 protocols
- 20: 11 < 12 protocols
- 21: 12 < 13 protocols
- 22: 13 < 14 protocols
- 23: 14 < 15 protocols
- 24: 15 < 16 protocols
- 25: 16 < 17 protocols

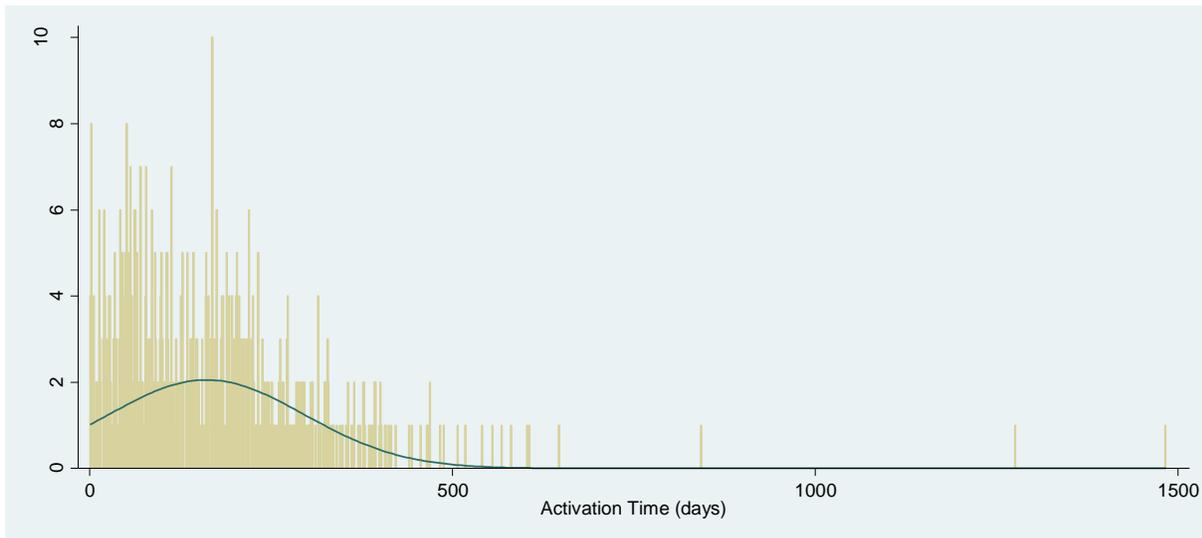
APPENDIX E: SITE-LEVEL STUDY CHARACTERISTIC DISTRIBUTION CHARTS

Appendix E-1: Distribution of Activation Time in Days

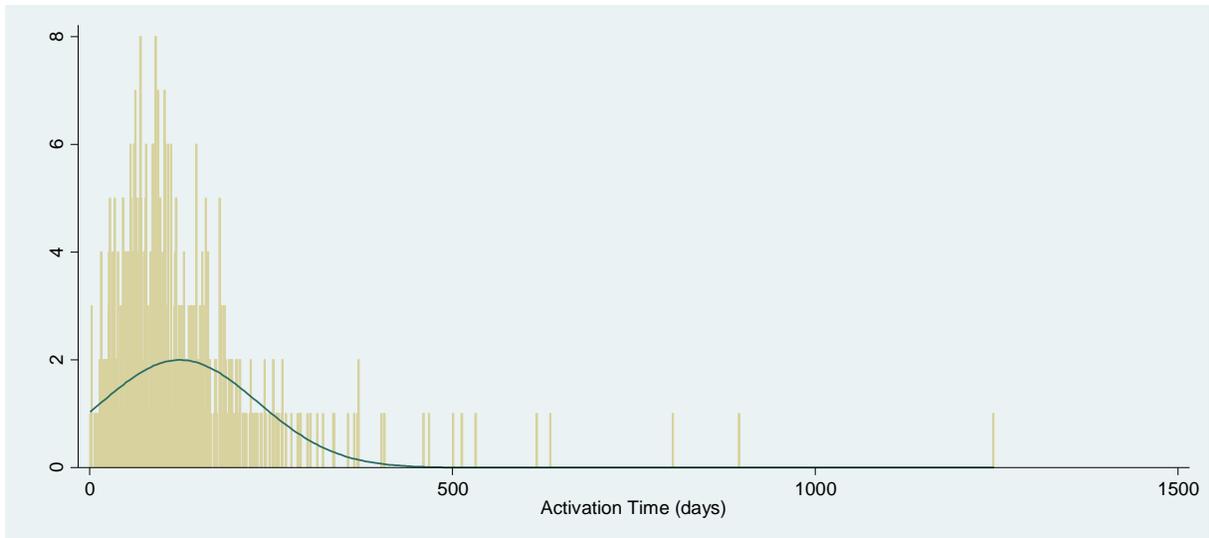
Appendix E-1a: Site 173472



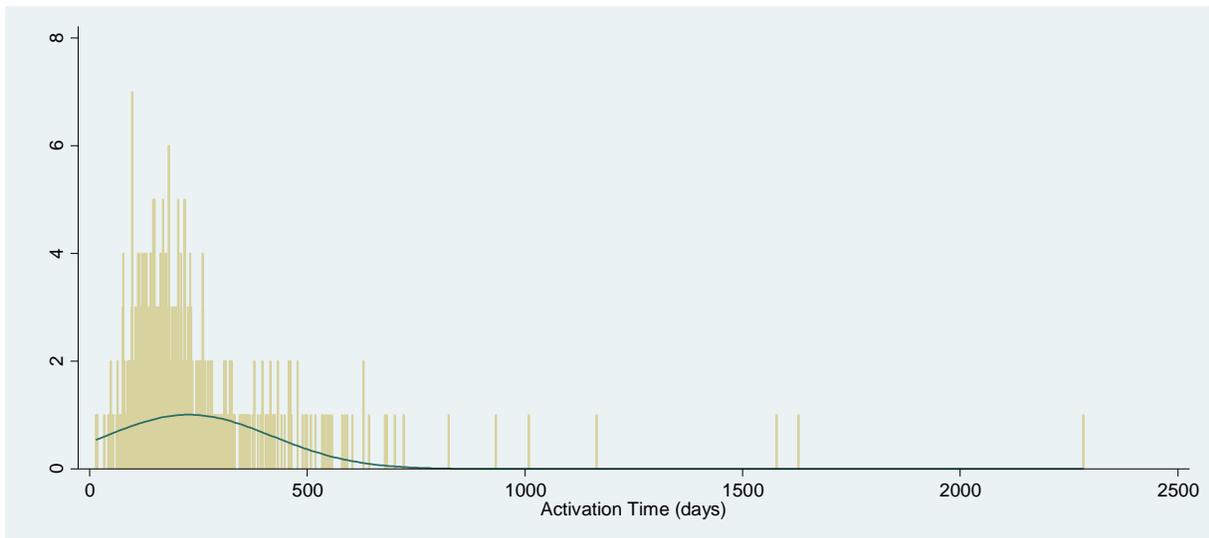
Appendix E-1b: Site 448155



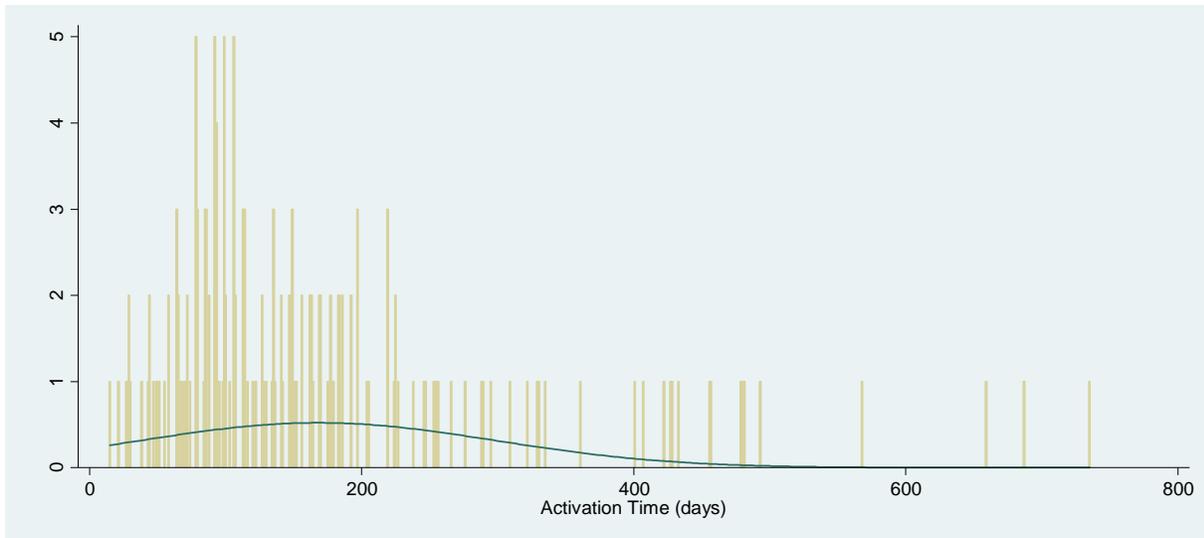
Appendix E-1c: Site 494048



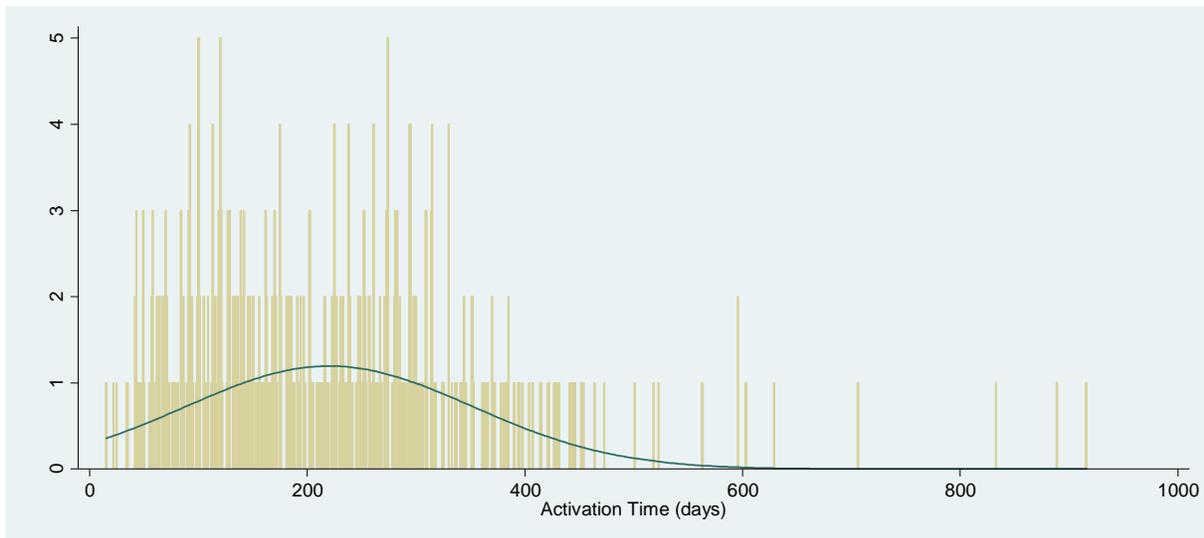
Appendix E-1d: Site 512786



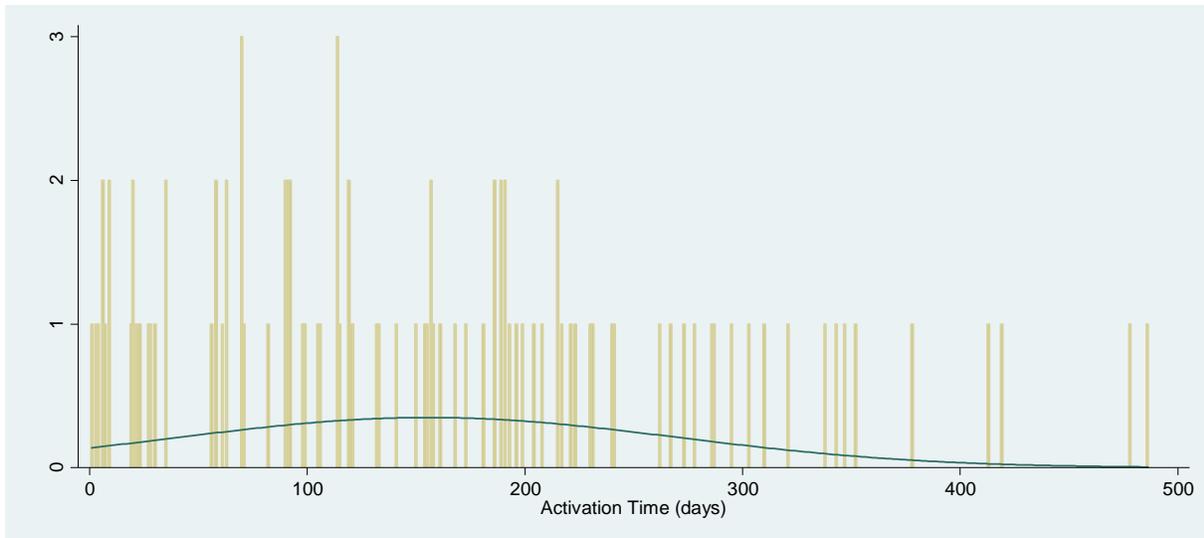
Appendix E-1e: Site 560623



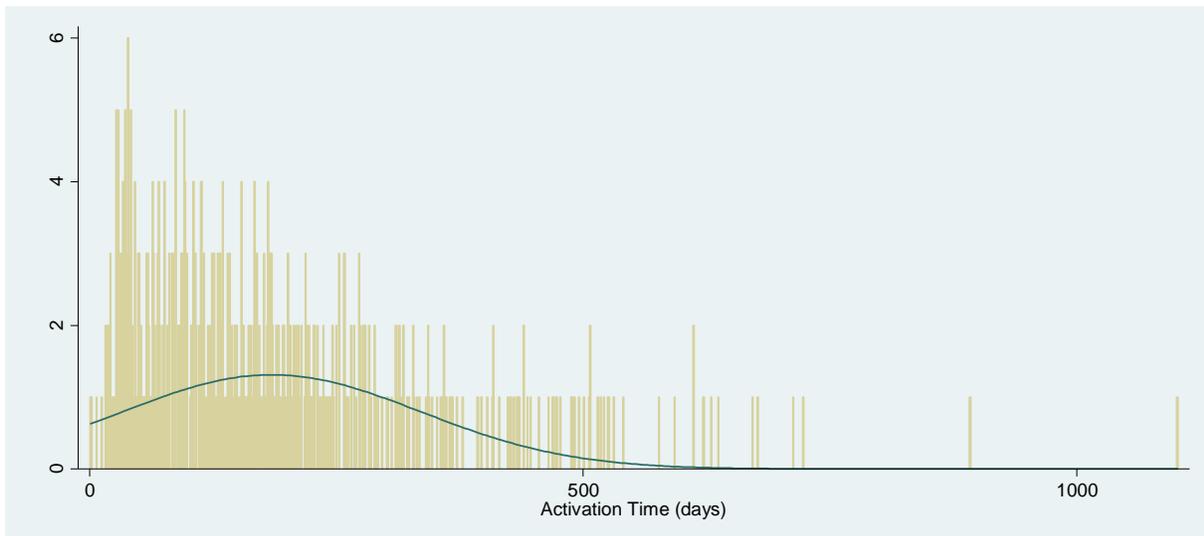
Appendix E-1f: Site 602591



Appendix E-1g: Site 696337

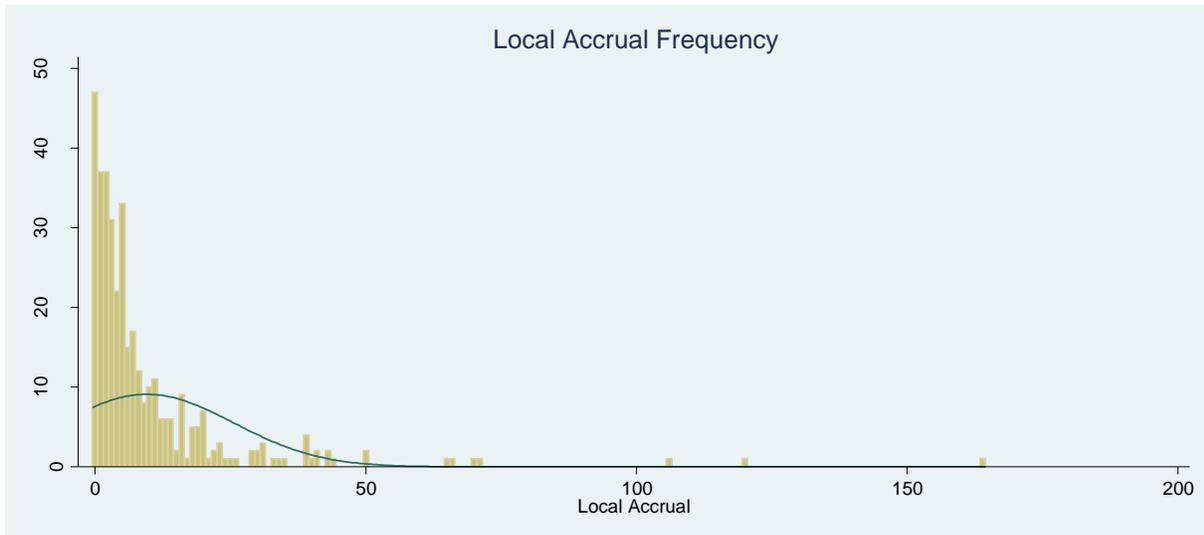


Appendix E-1h: Site 714145

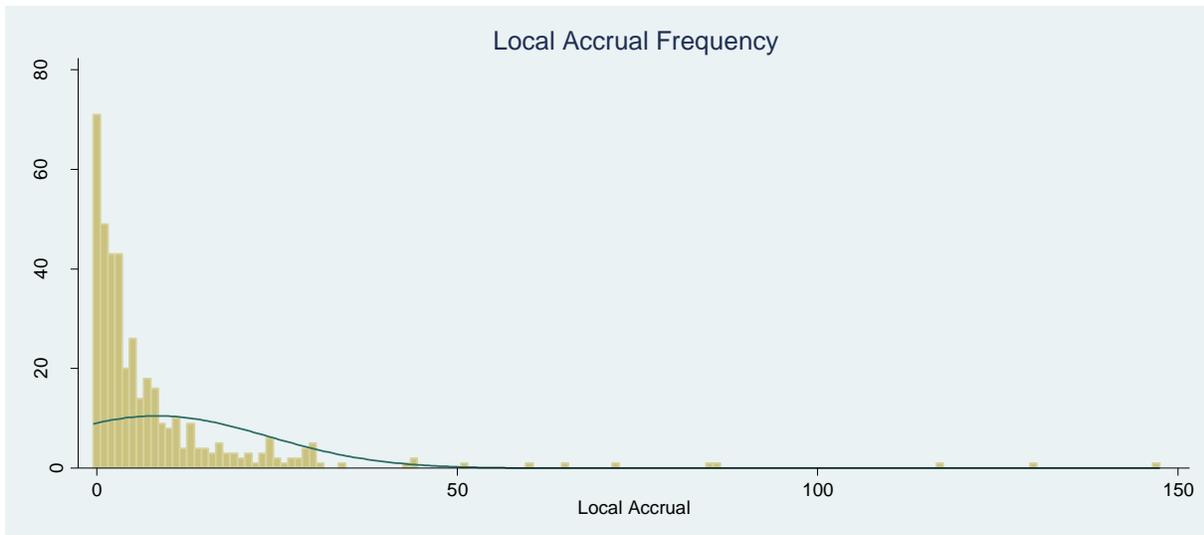


Appendix E-2: Distribution of Actual Accrual

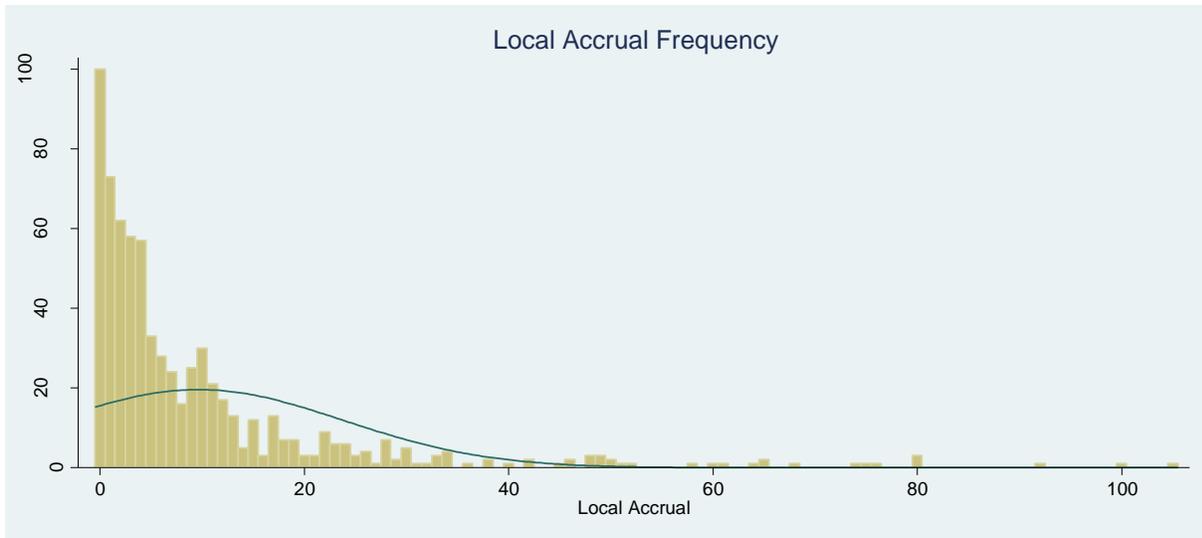
Appendix E-2a: Site 104647



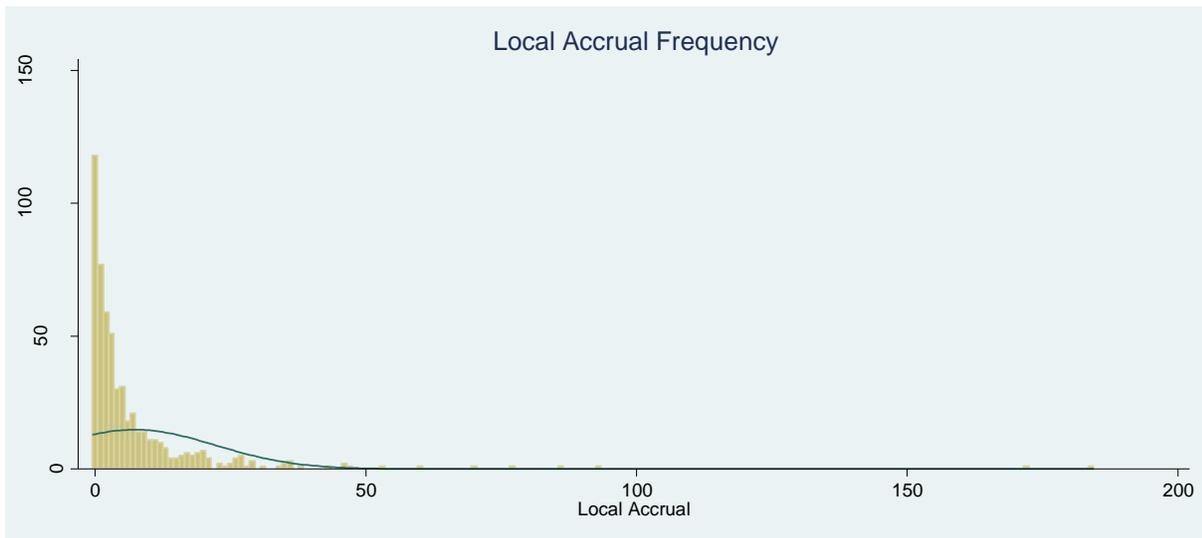
Appendix E-2b: Site 173472



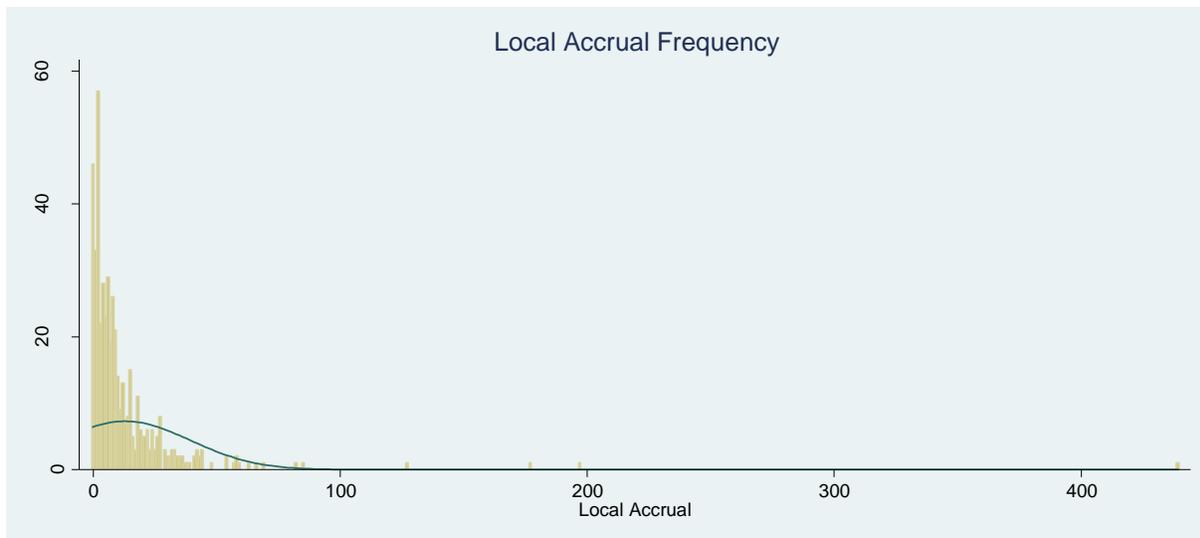
Appendix E-2c: Site 448155



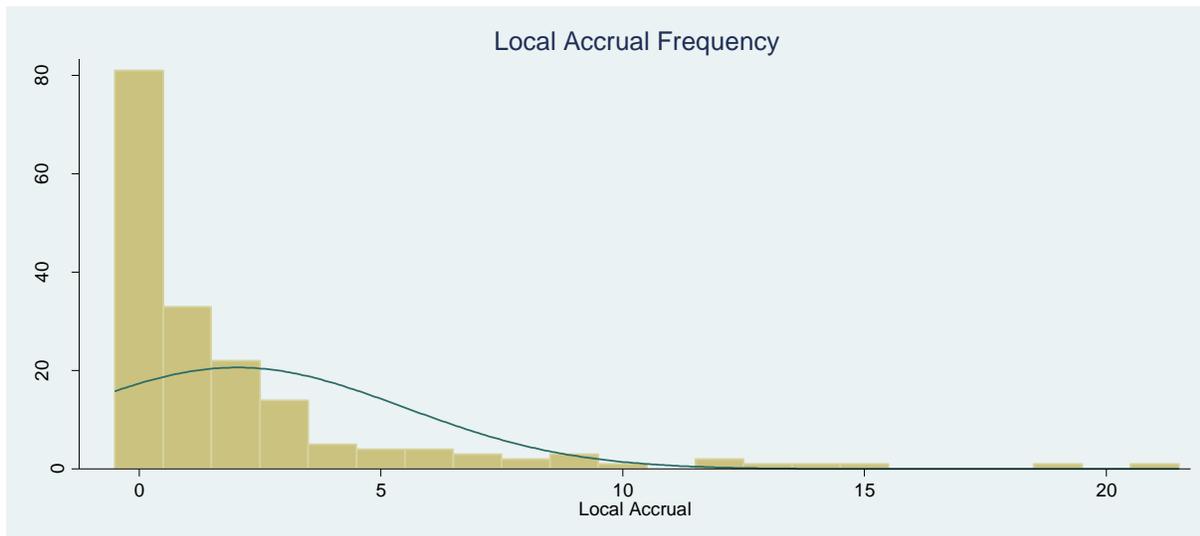
Appendix E-2d: Site 494048



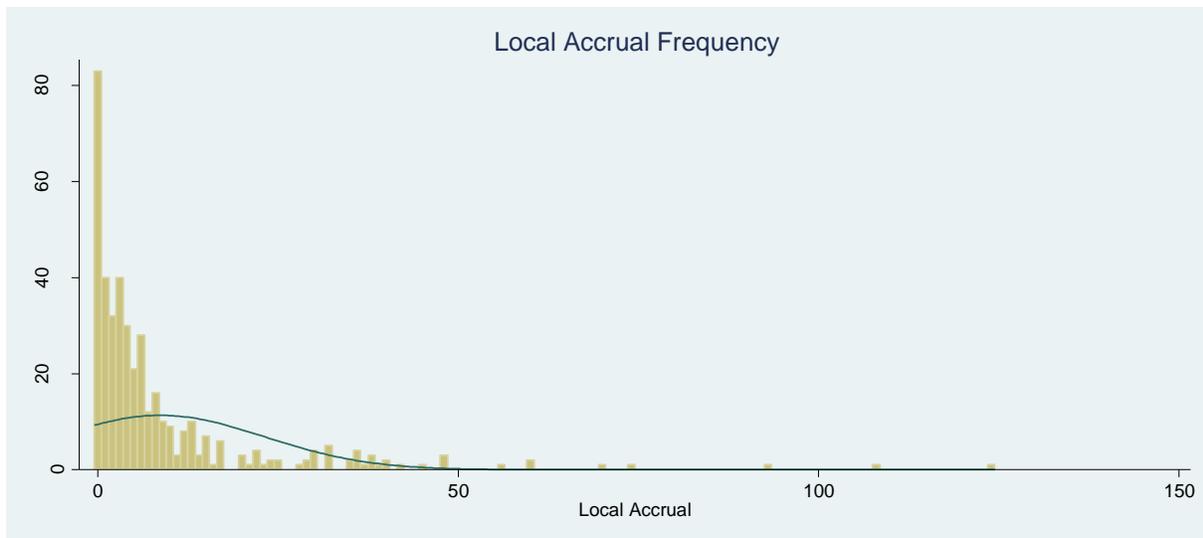
Appendix E-2e: Site 512786



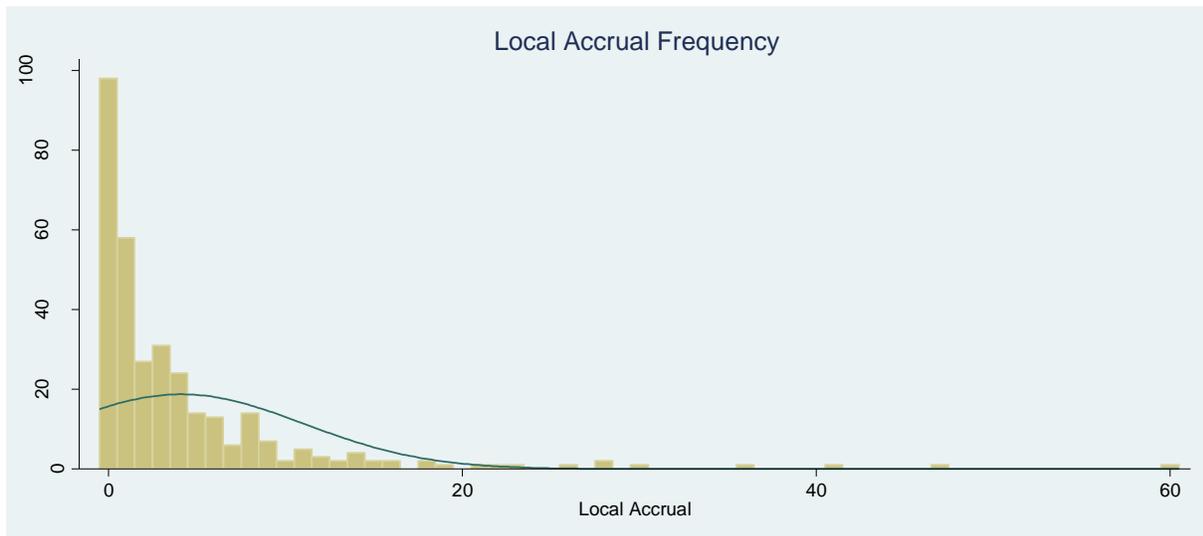
Appendix E-2f: Site 560623



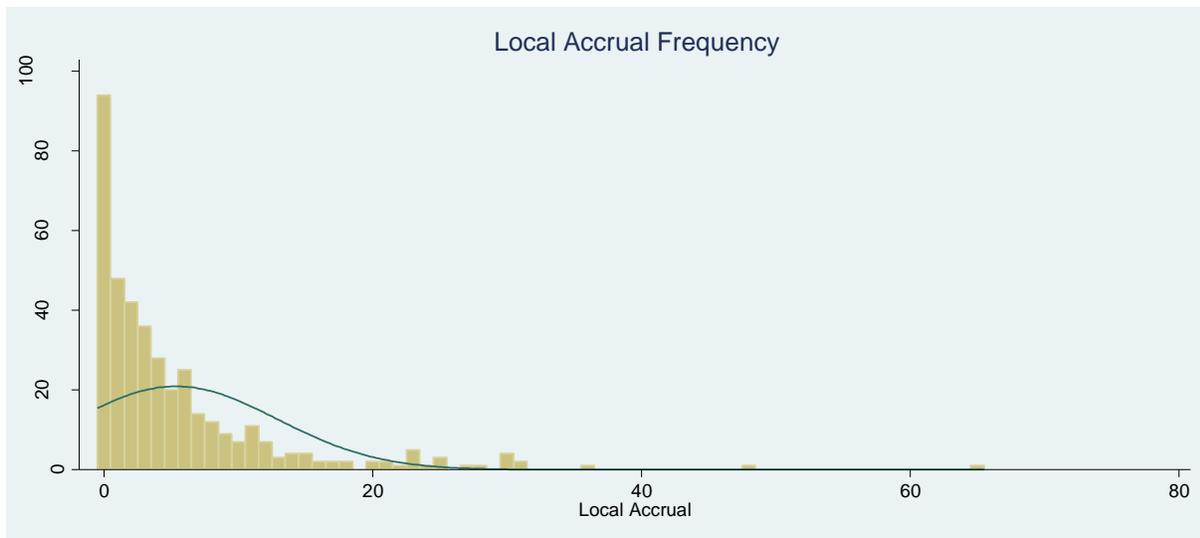
Appendix E-2g: Site 575415



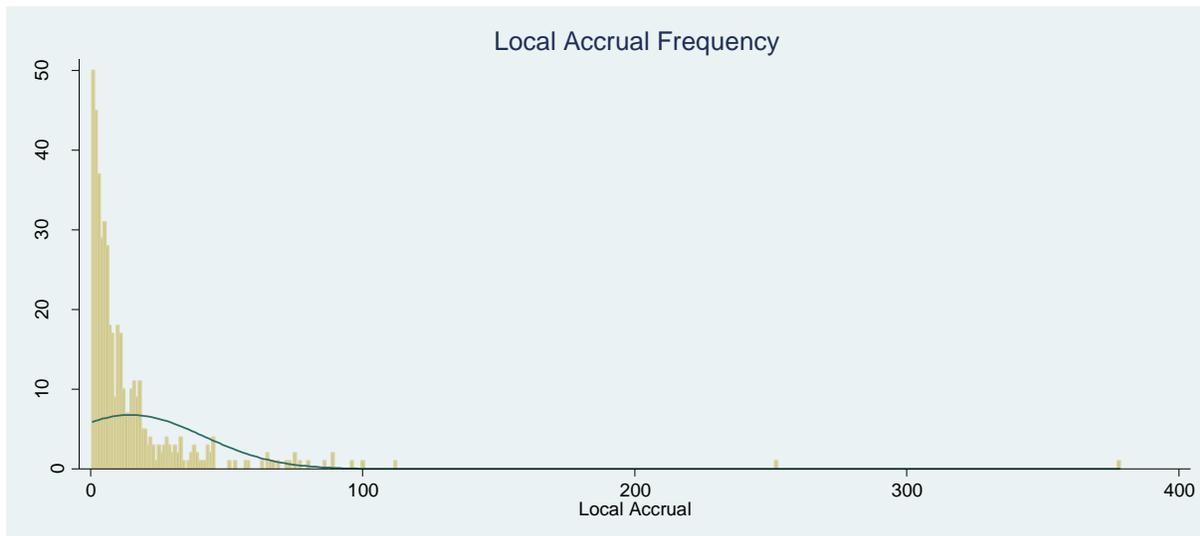
Appendix E-2h: Site 598430



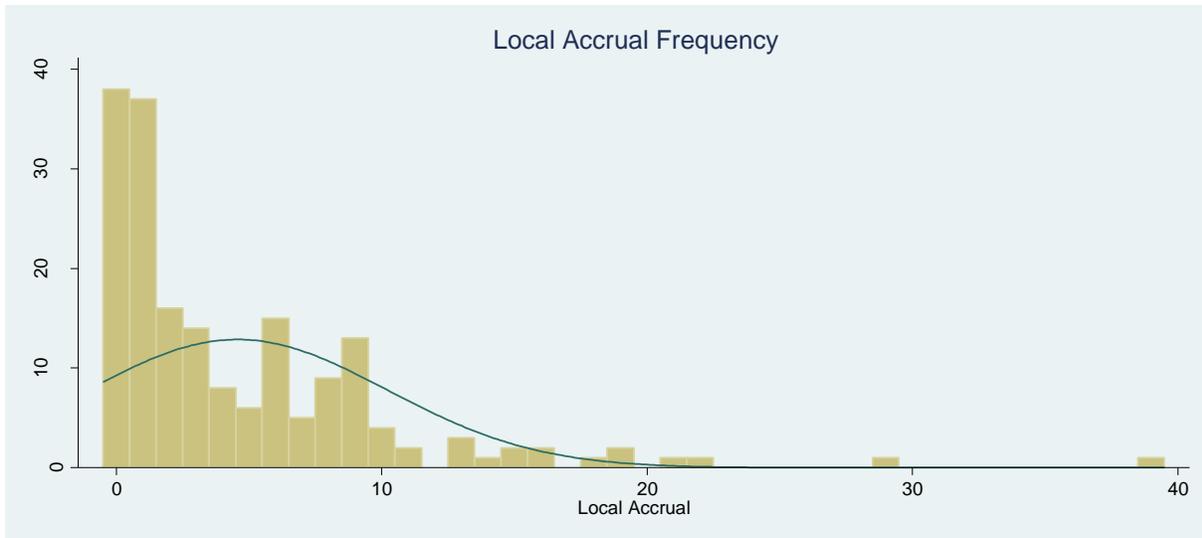
Appendix E-2i: Site 602591



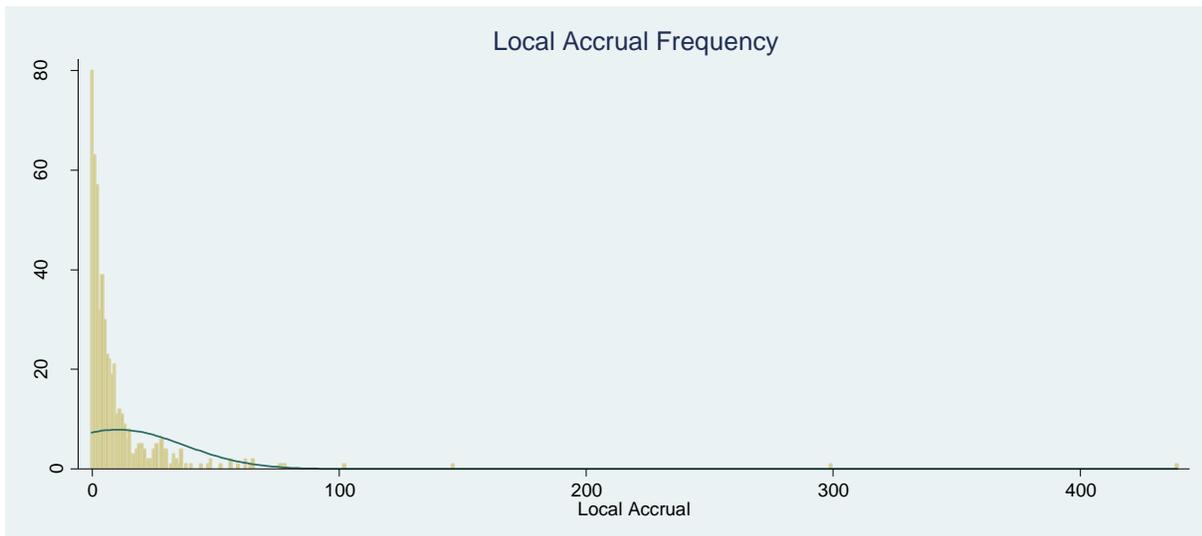
Appendix E-2j: Site 689326



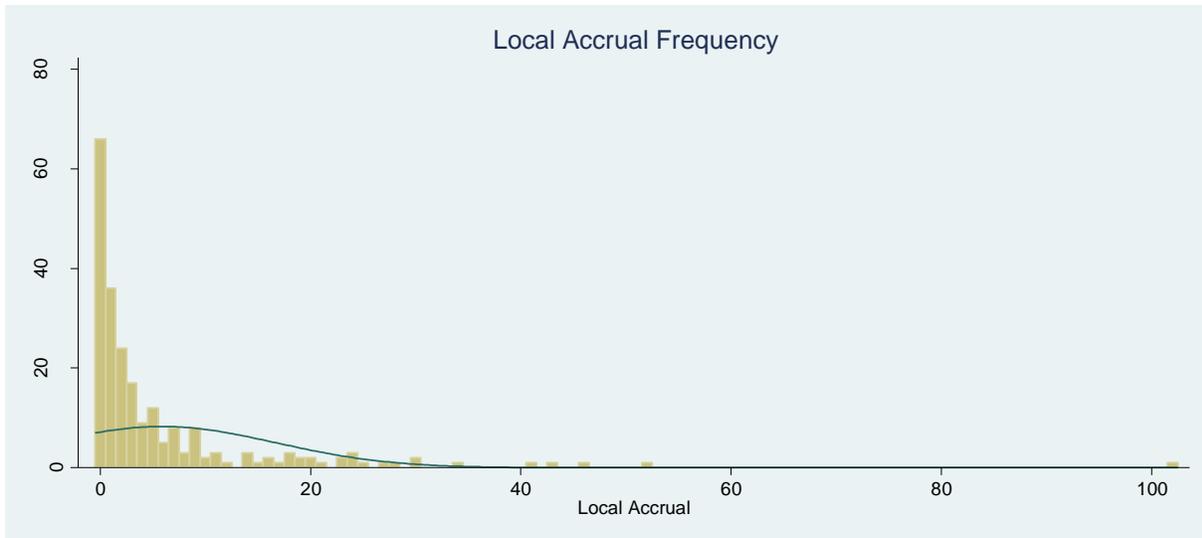
Appendix E-2k: Site 696337



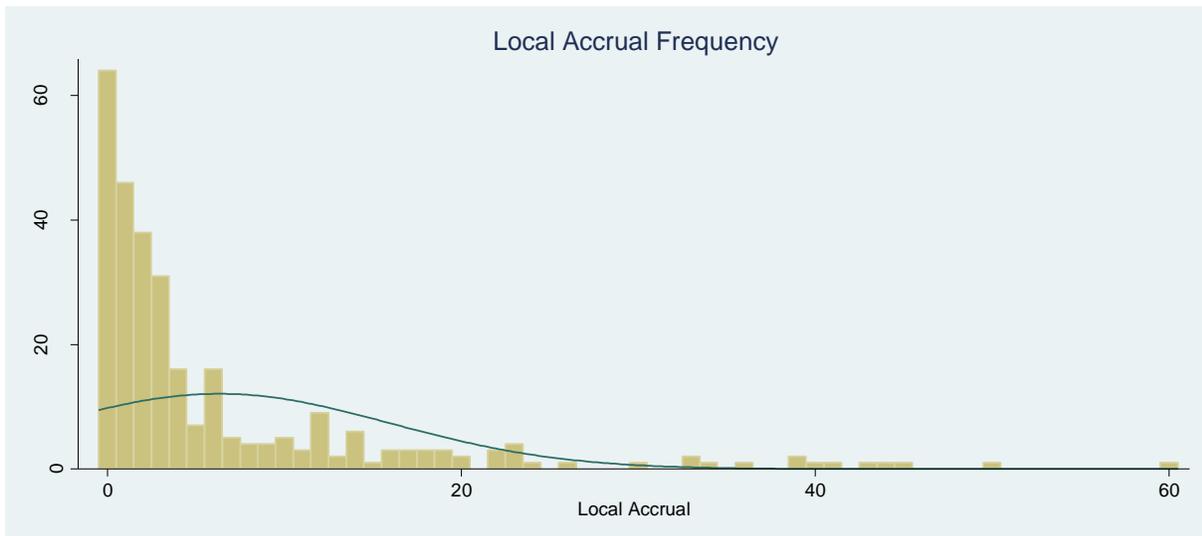
Appendix E-2l: Site 714145



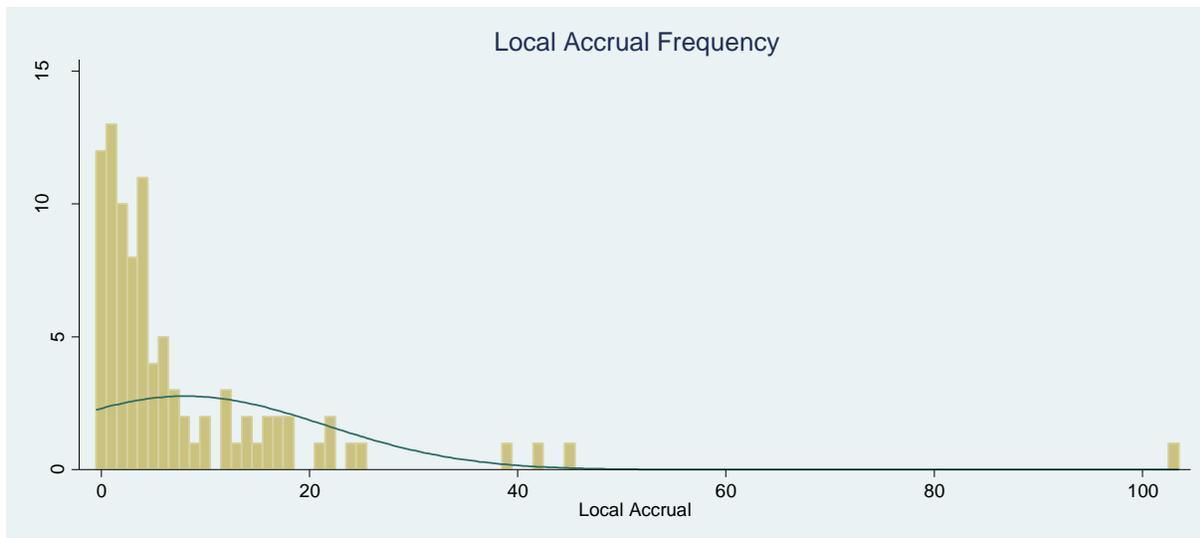
Appendix E-2m: Site 715532



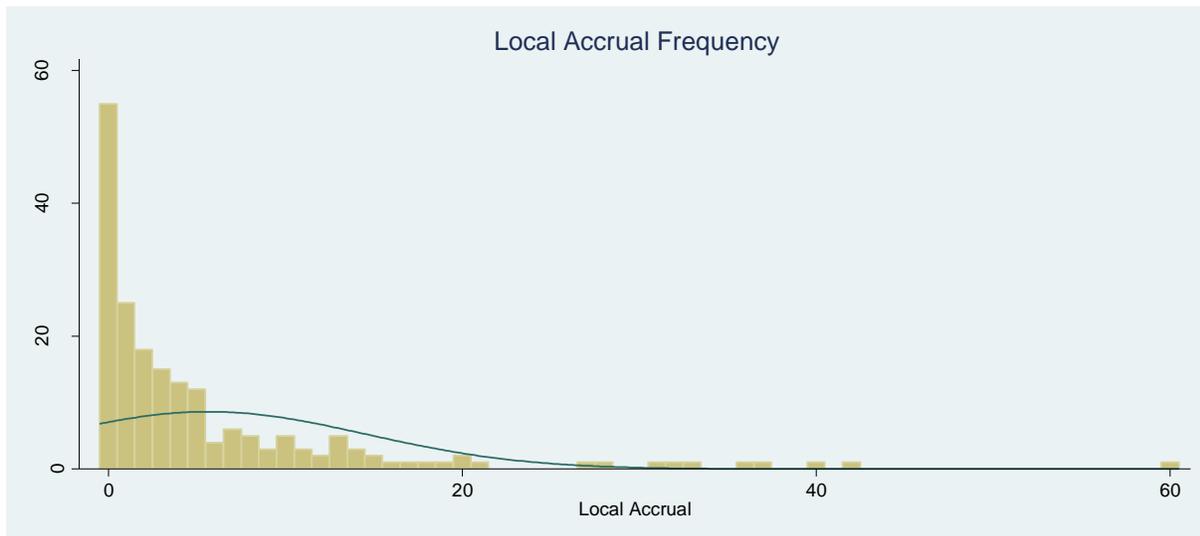
Appendix E-2n: Site 846594



Appendix E-2o: Site 997056



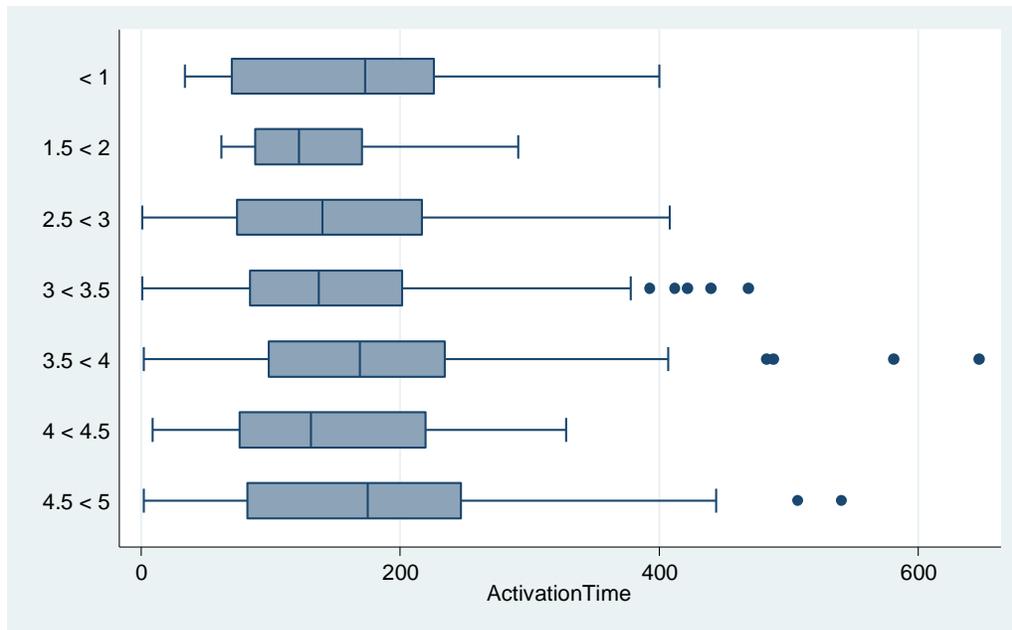
Appendix E-2p: Site 998666



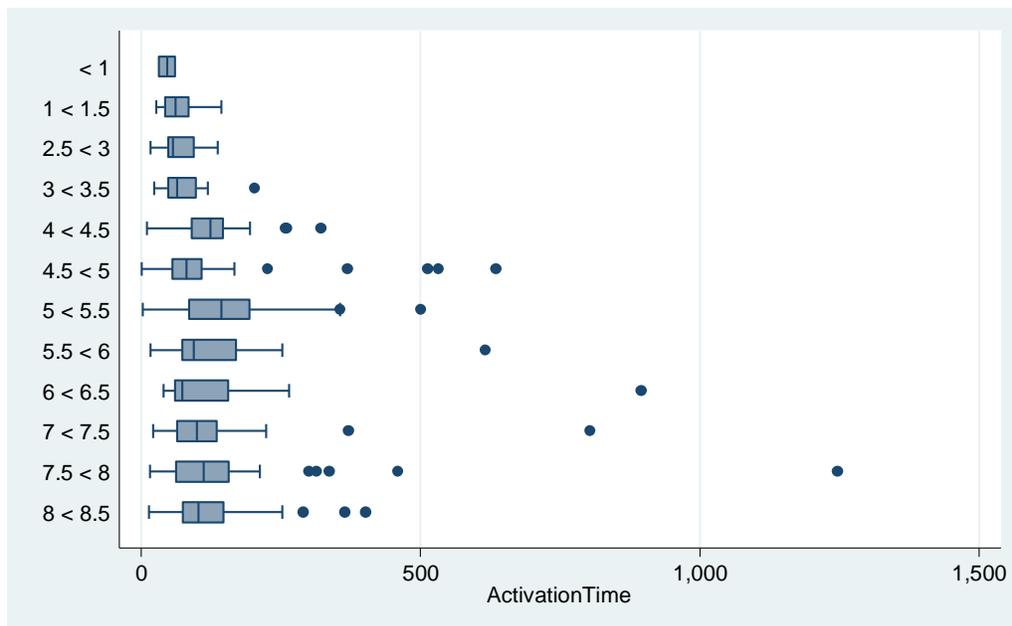
APPENDIX F: SPECIFIC AIM TWO – SITE-SPECIFIC PROTOCOL WORKLOAD MODELS

Appendix F-1: Protocols per Full-Time Equivalent (FTE) versus Time to Study Activation

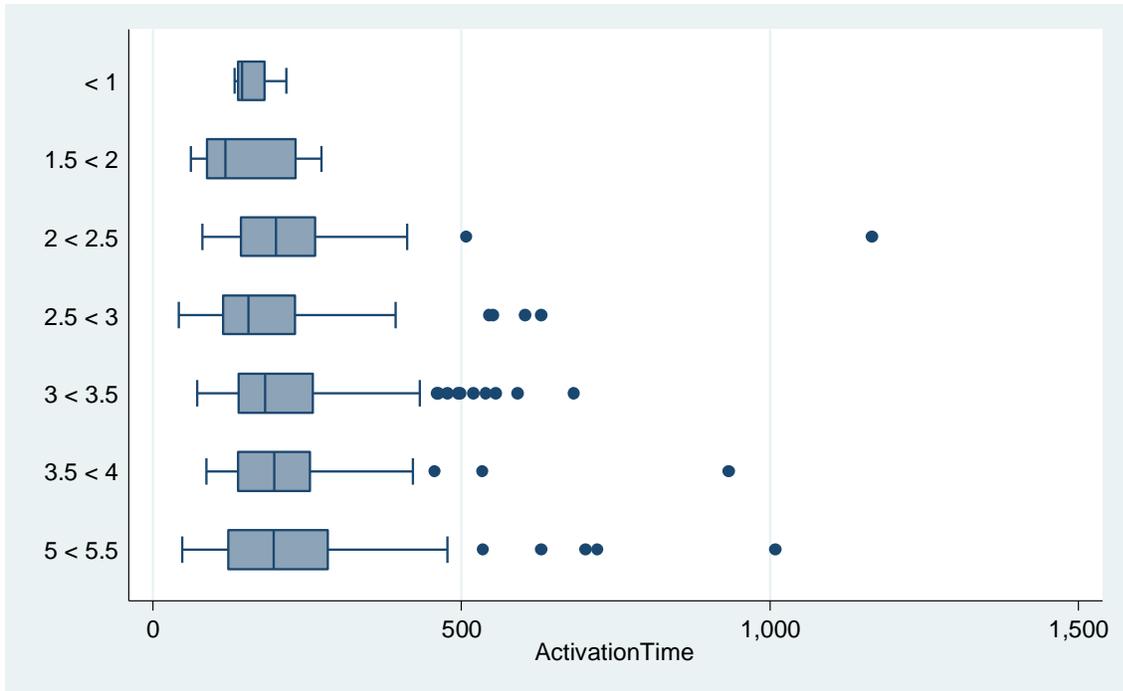
Appendix F-1a: Site 448155



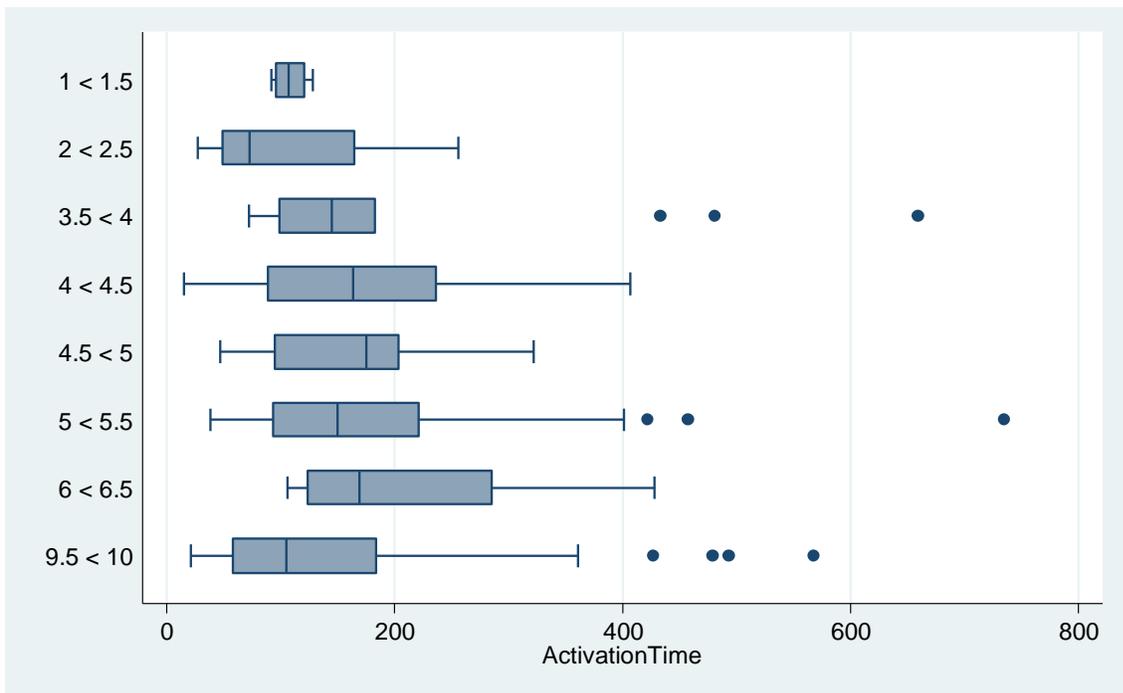
Appendix F-1b: Site 494048



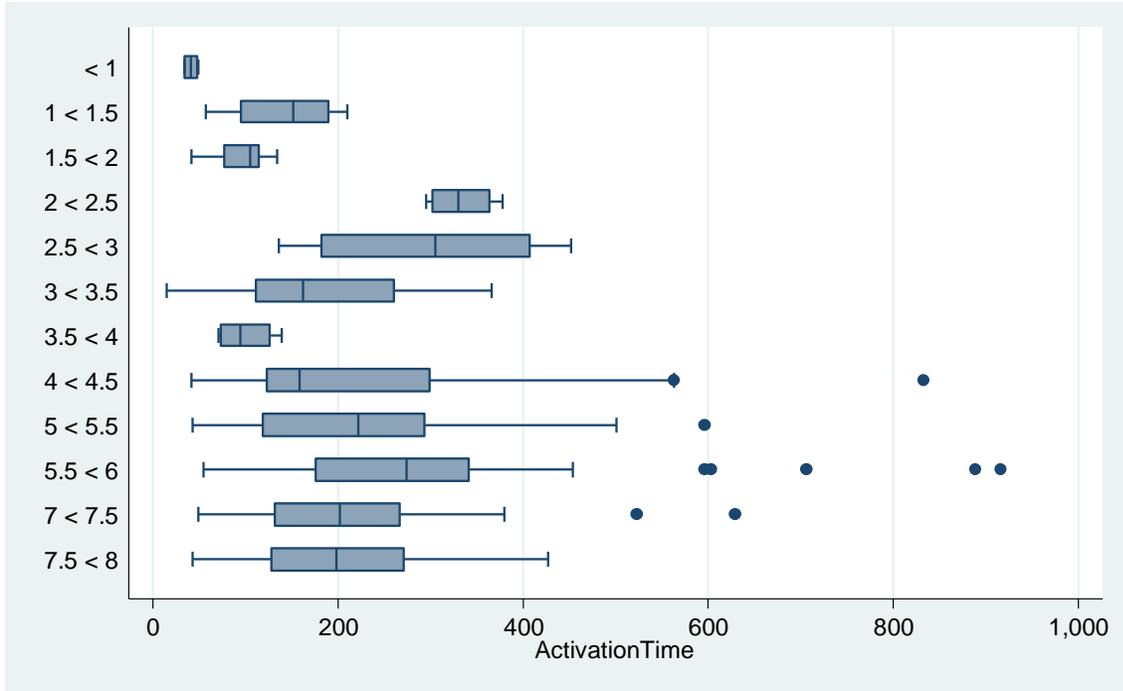
Appendix F-1c: Site 512786



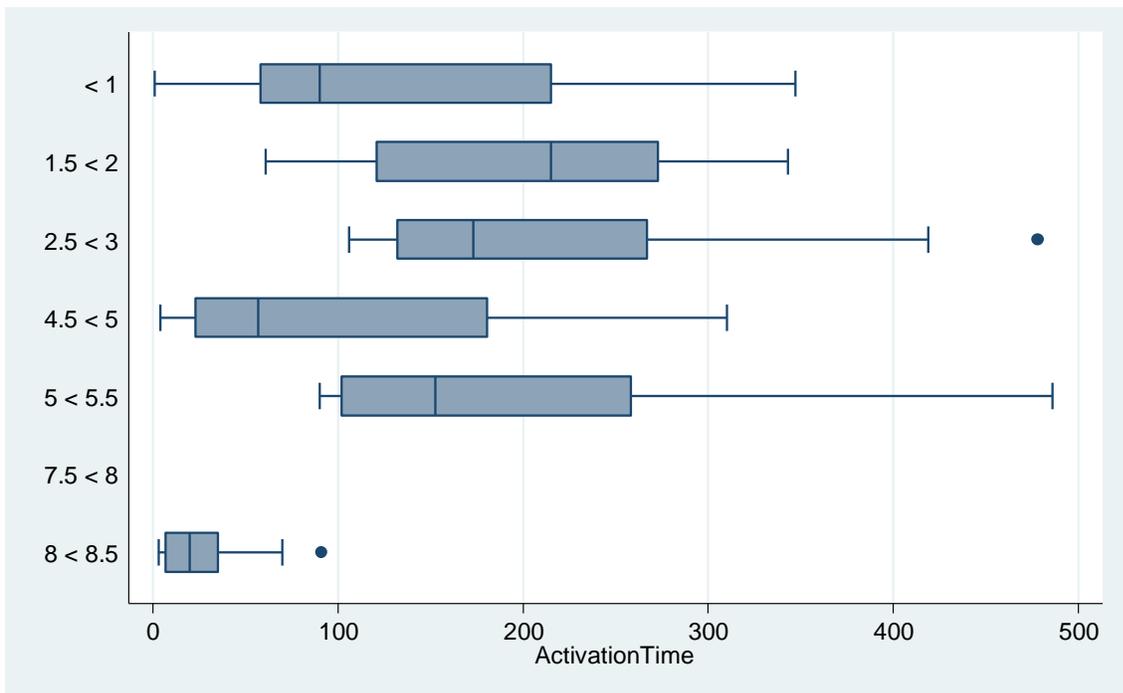
Appendix F-1d: Site 560623



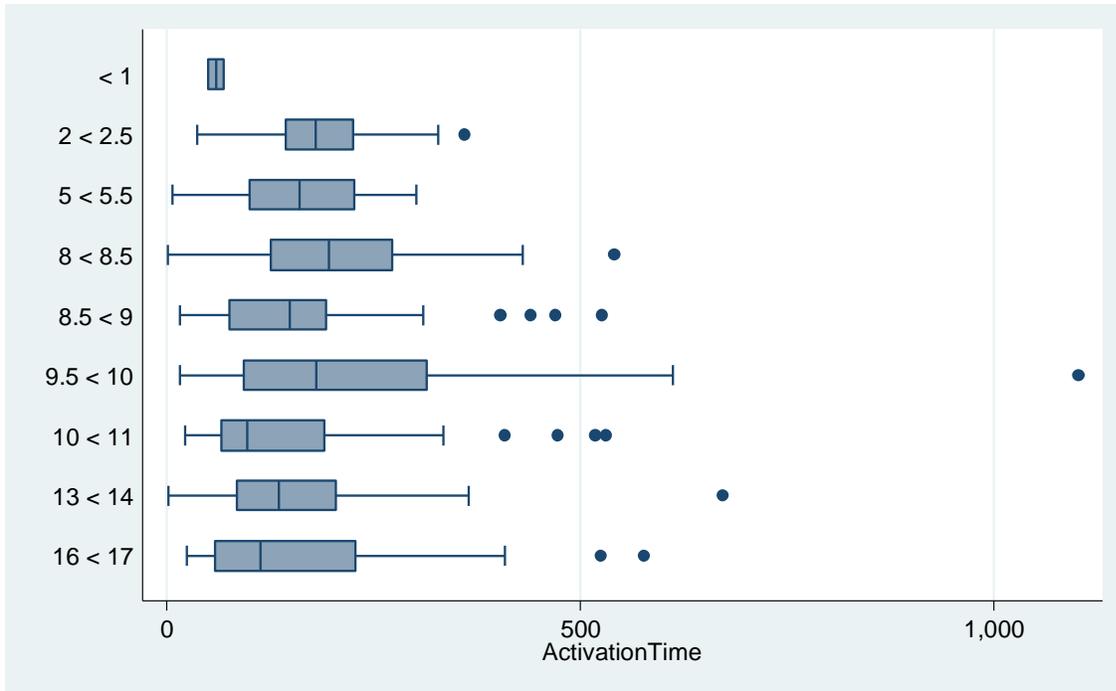
Appendix F-1e: Site 602591



Appendix F-1f: Site 696337

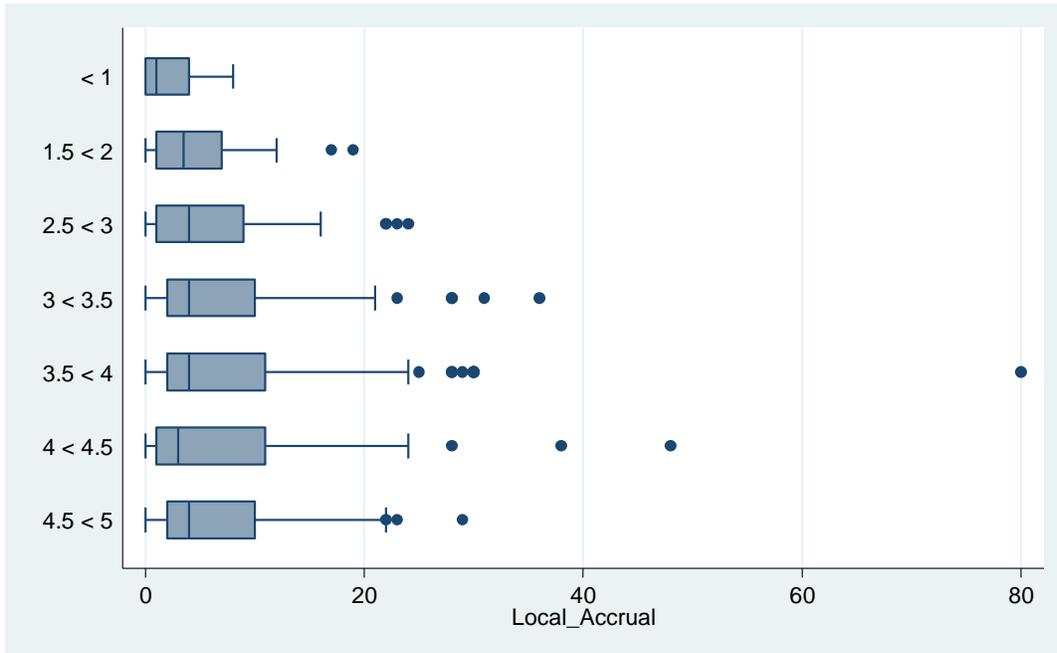


Appendix F-1g: Site 714145

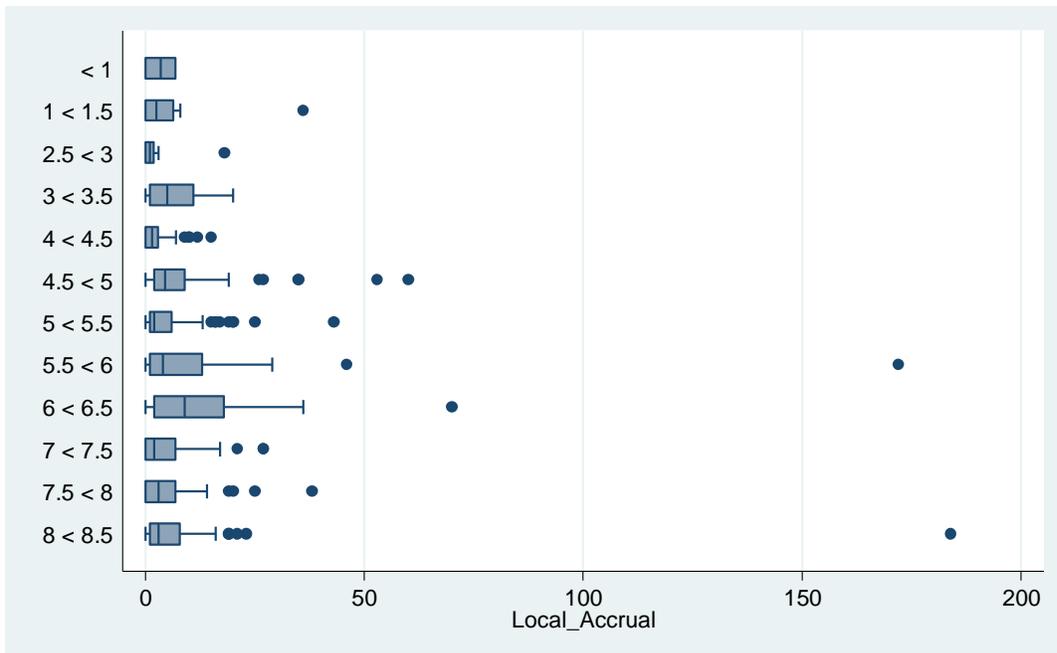


Appendix F-2: Protocols per Full-Time Equivalent versus Overall Protocol Accrual

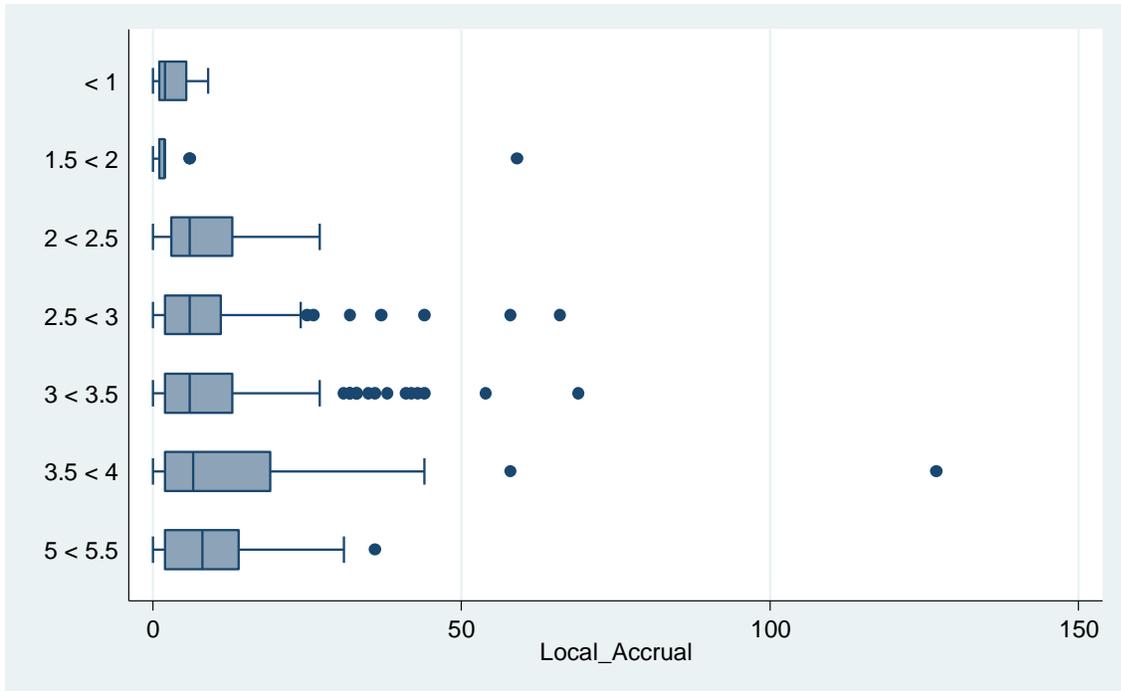
Appendix F-2a: Site 448155



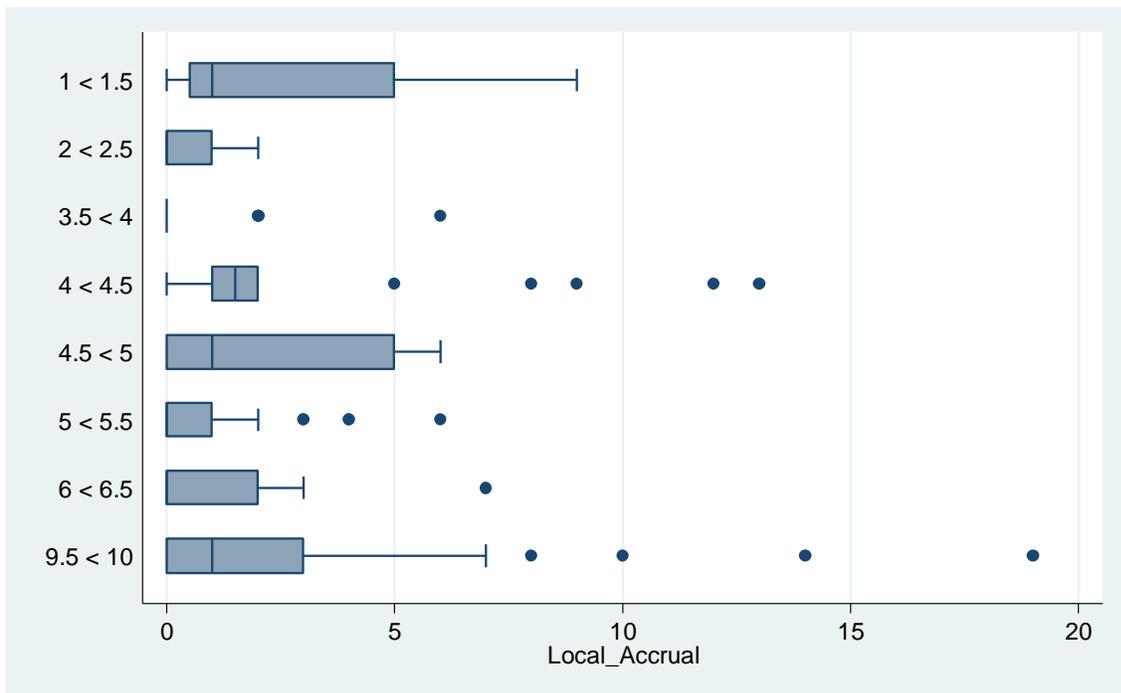
Appendix F-2b: Site 494048



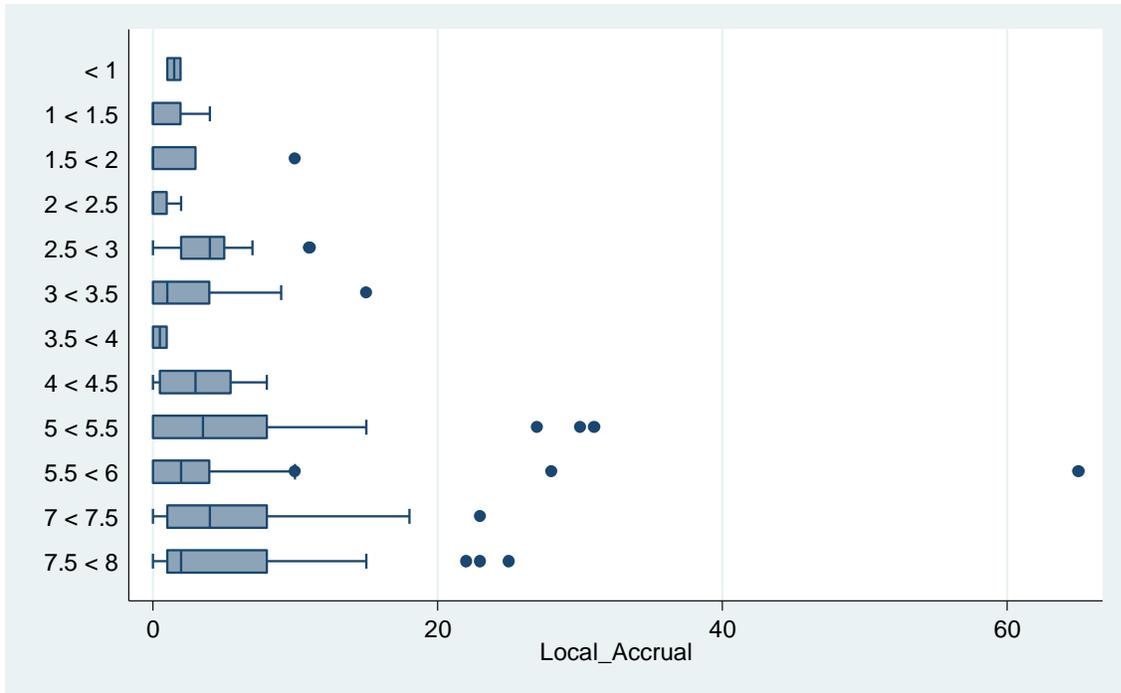
Appendix F-2c: Site 512786



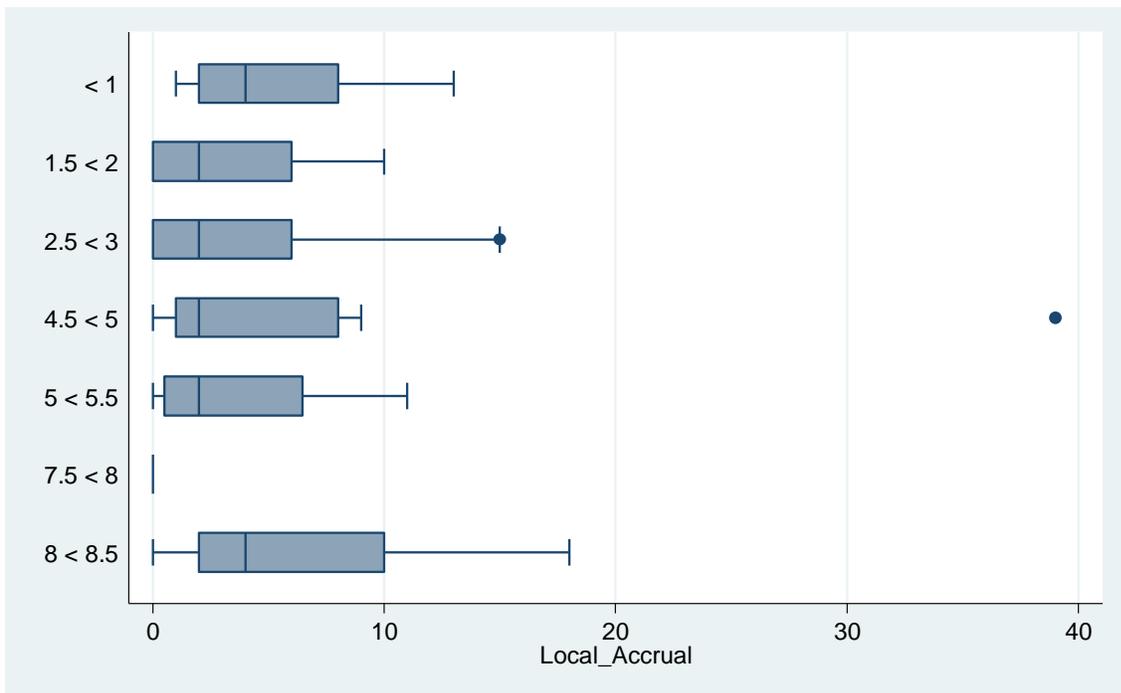
Appendix F-2d: Site 560623



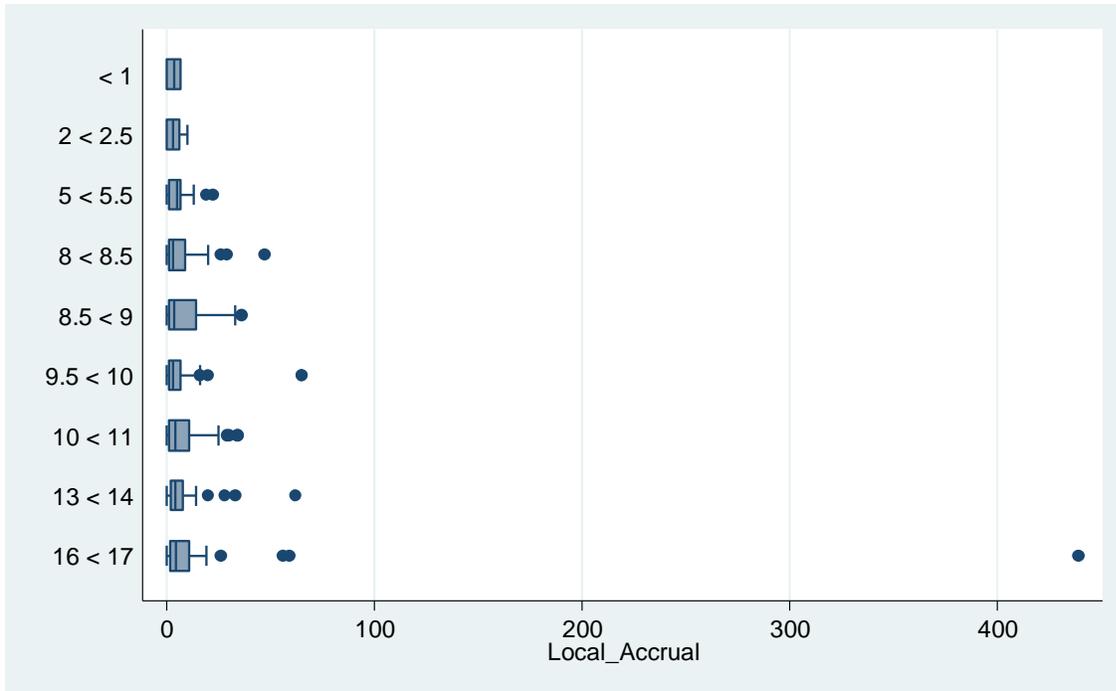
Appendix F-2e: Site 602591



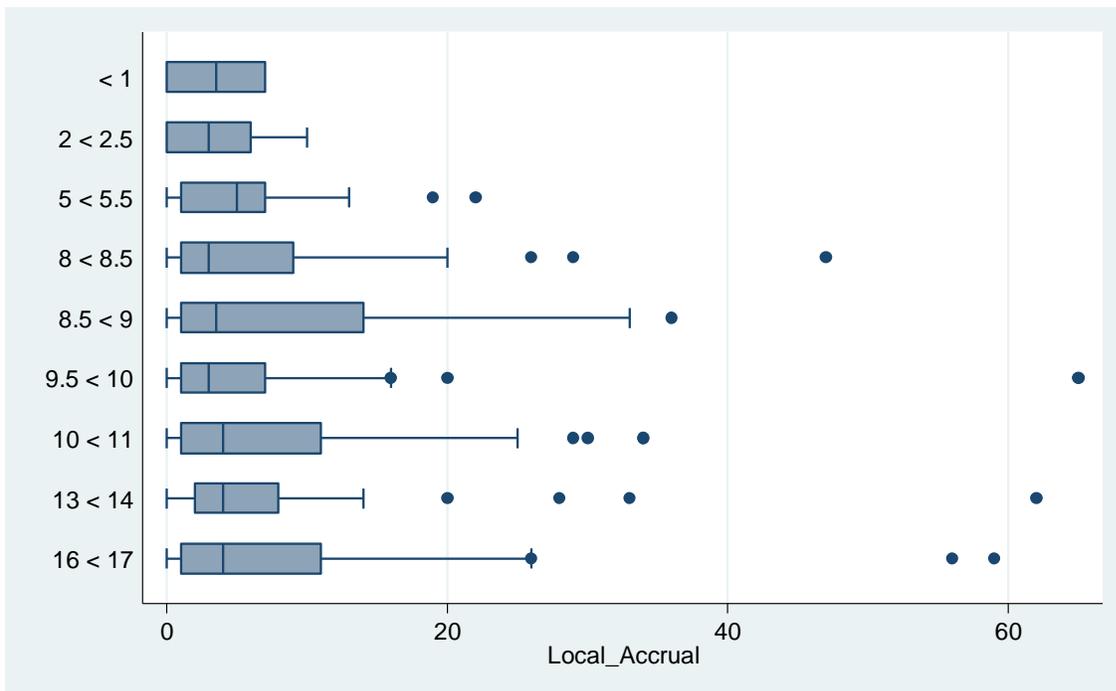
Appendix F-2f: Site 696337



Appendix F-2g: Site 714145

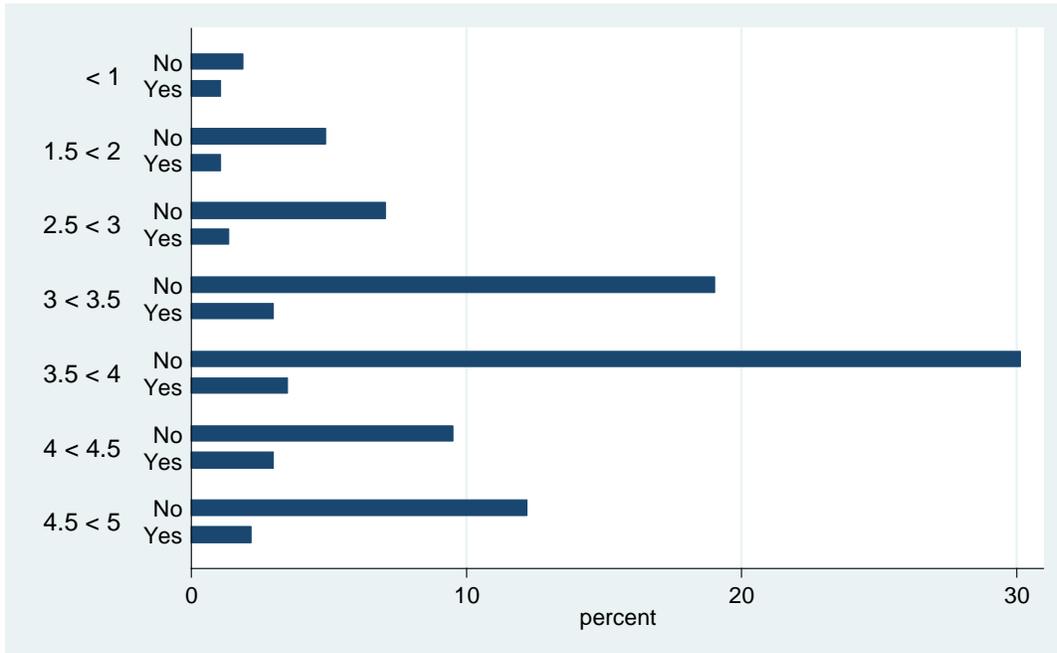


Appendix F-2h: Site 714145 – Under 100 Subjects

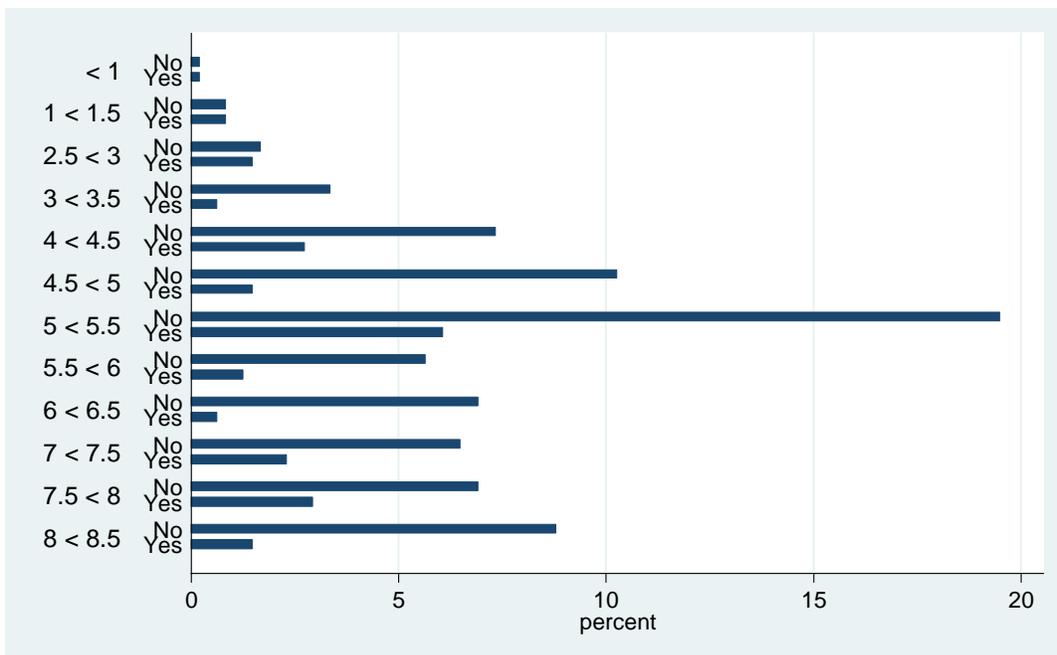


Appendix F-3: Protocols per Full-Time Equivalent by Zero-Accruing Status

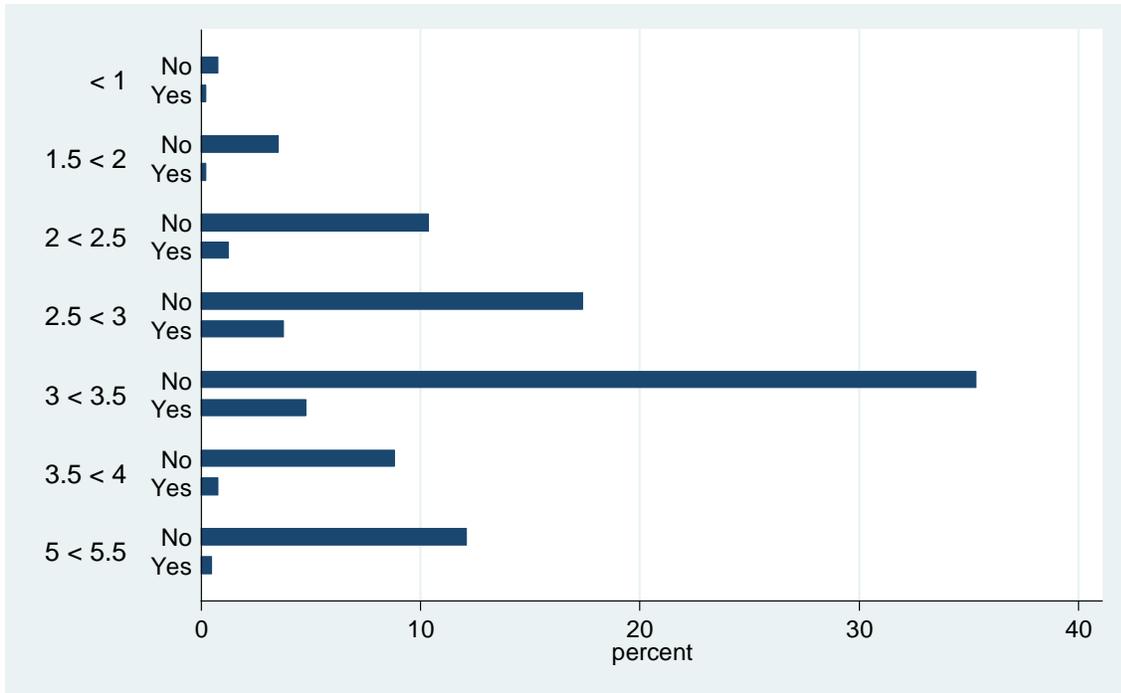
Appendix F-3a: Site 448155



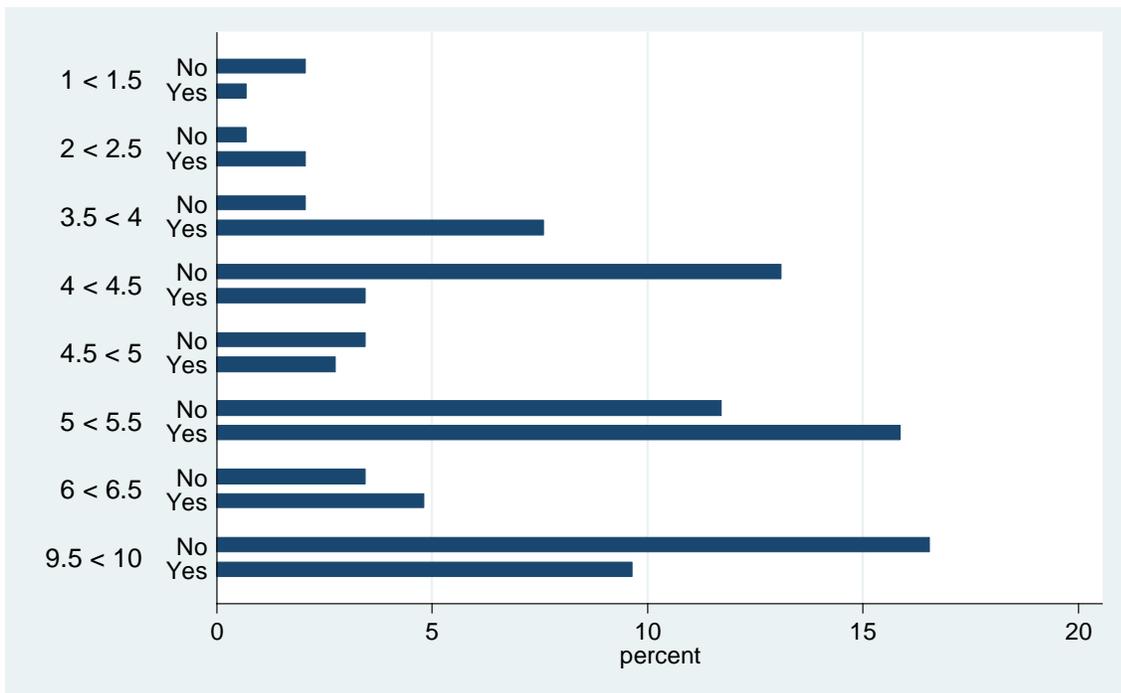
Appendix F-3b: Site 494048



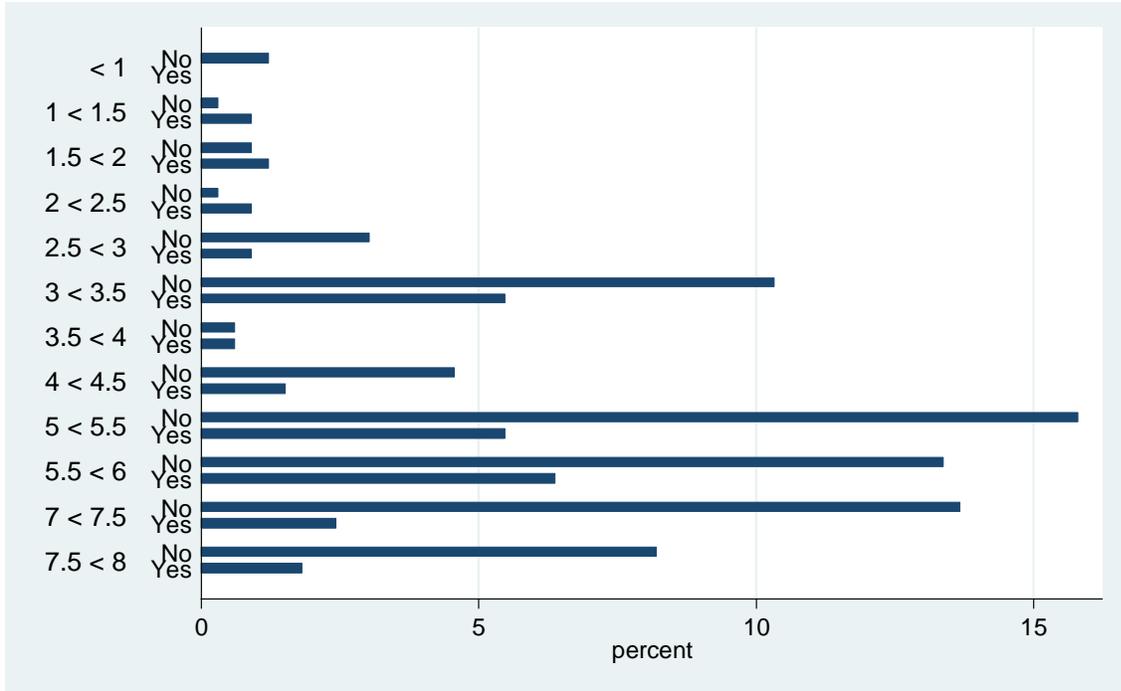
Appendix F-3c: Site 512786



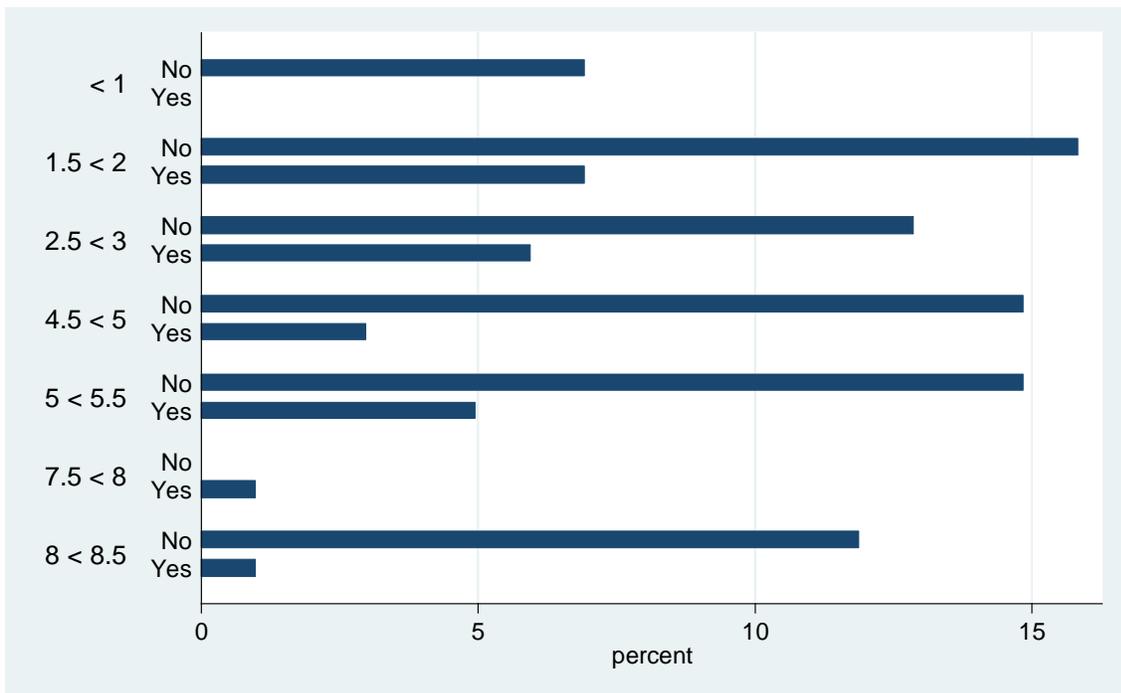
Appendix F-3d: Site 560623



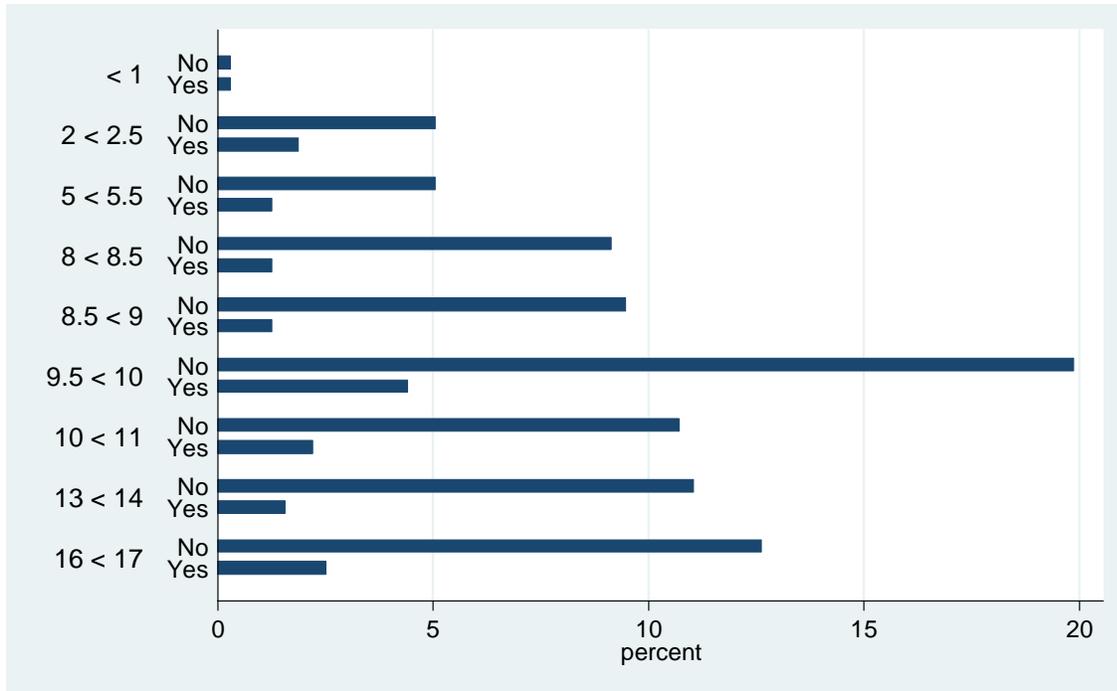
Appendix F-3e: Site 602591



Appendix F-3f: Site 696337

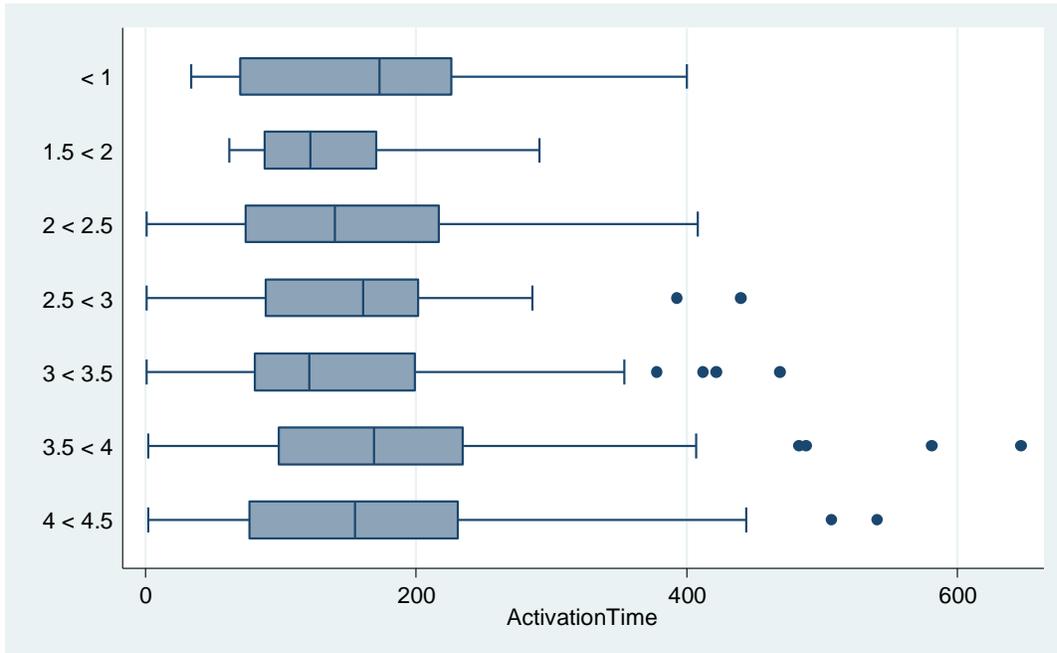


Appendix F-3g: Site 714145

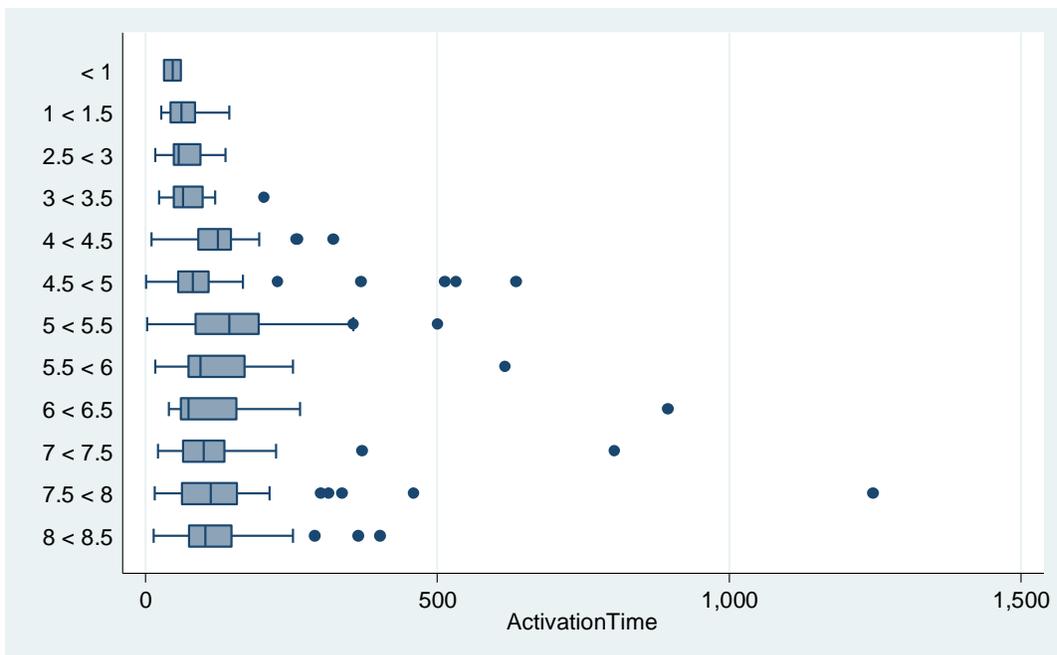


Appendix F-4: Protocols per Staff Member versus Time to Study Activation

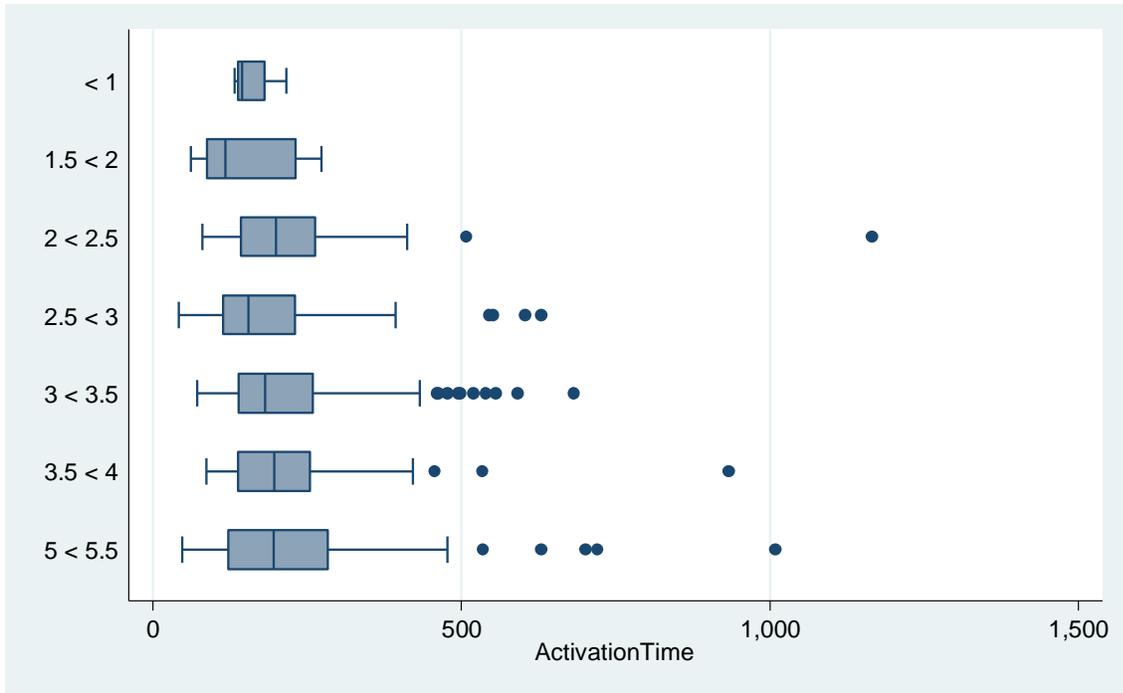
Appendix F-4a: Site 448155



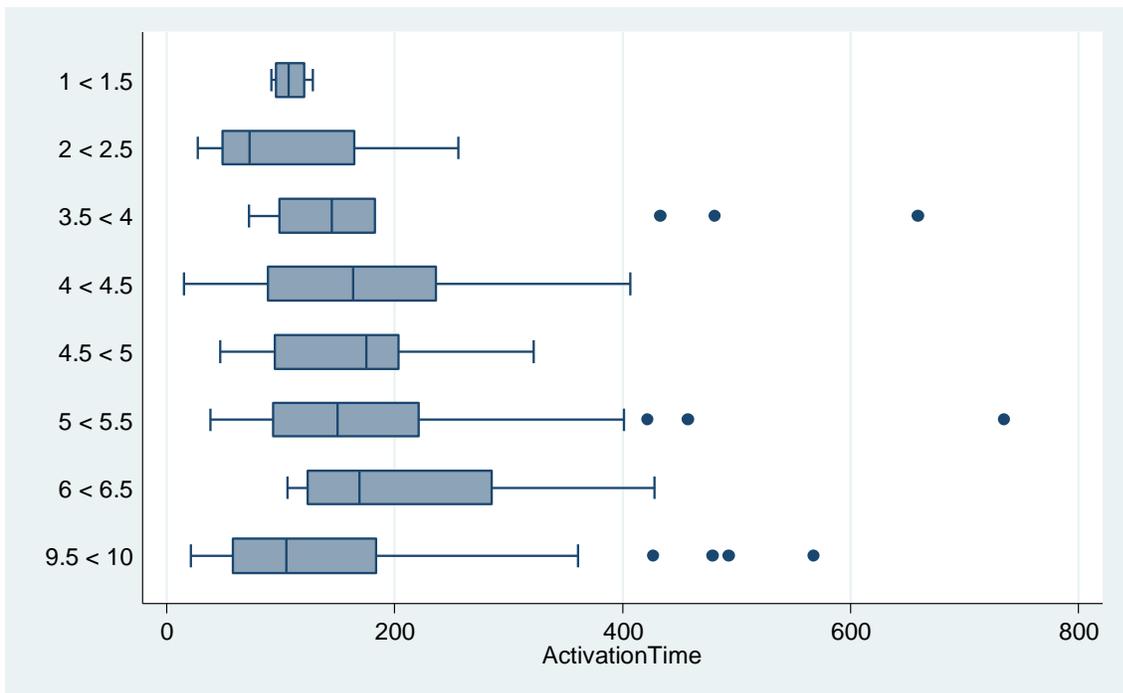
Appendix F-4b: Site 494048



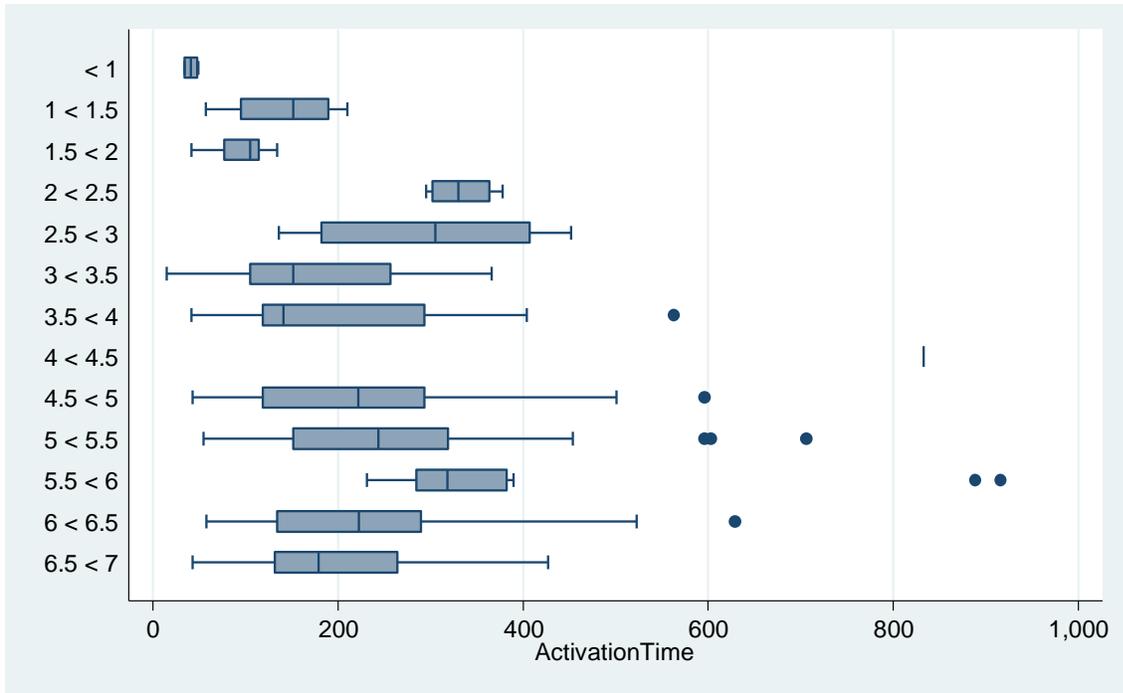
Appendix F-4c: Site 512786



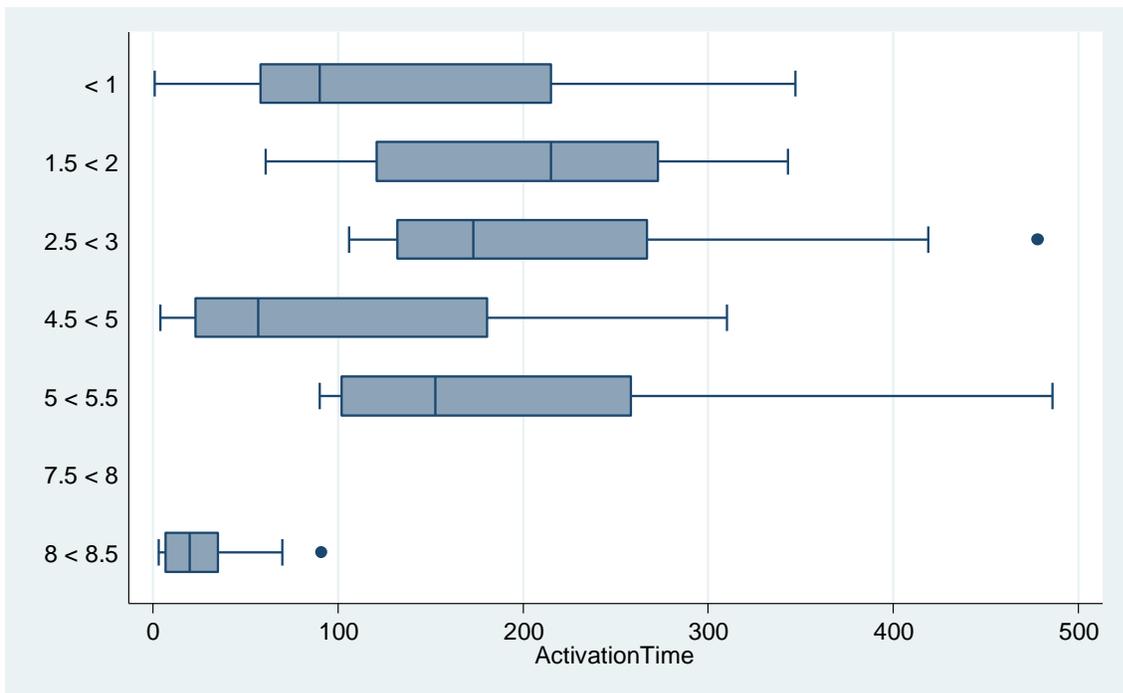
Appendix F-4d: Site 560623



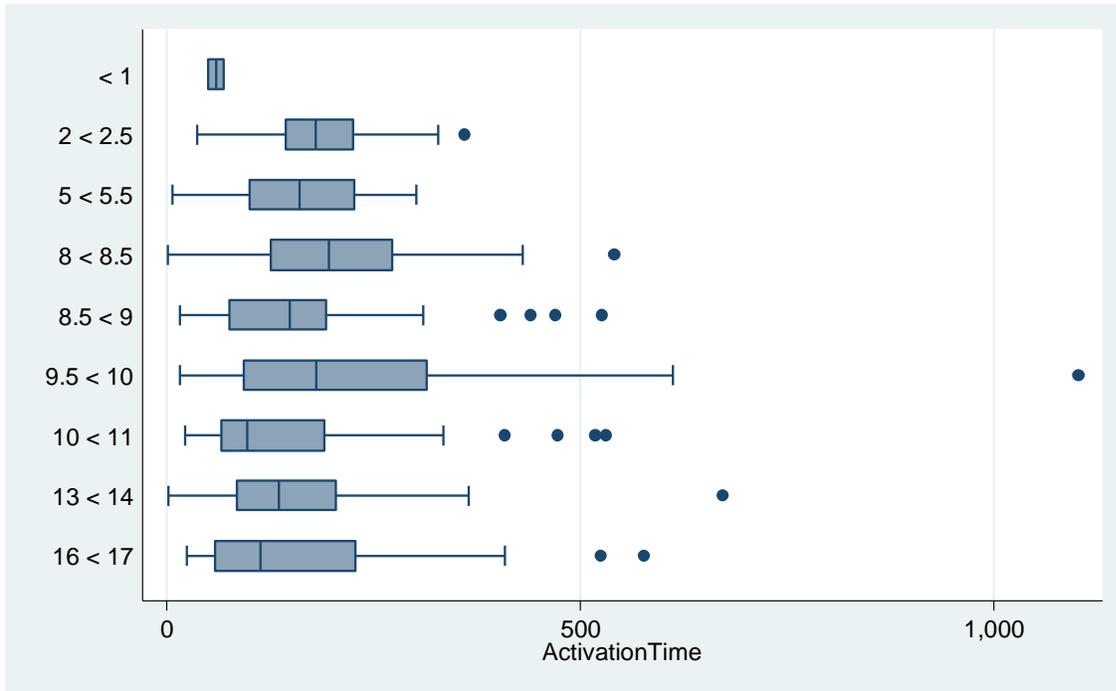
Appendix F-4e: Site 602591



Appendix F-4f: Site 696337

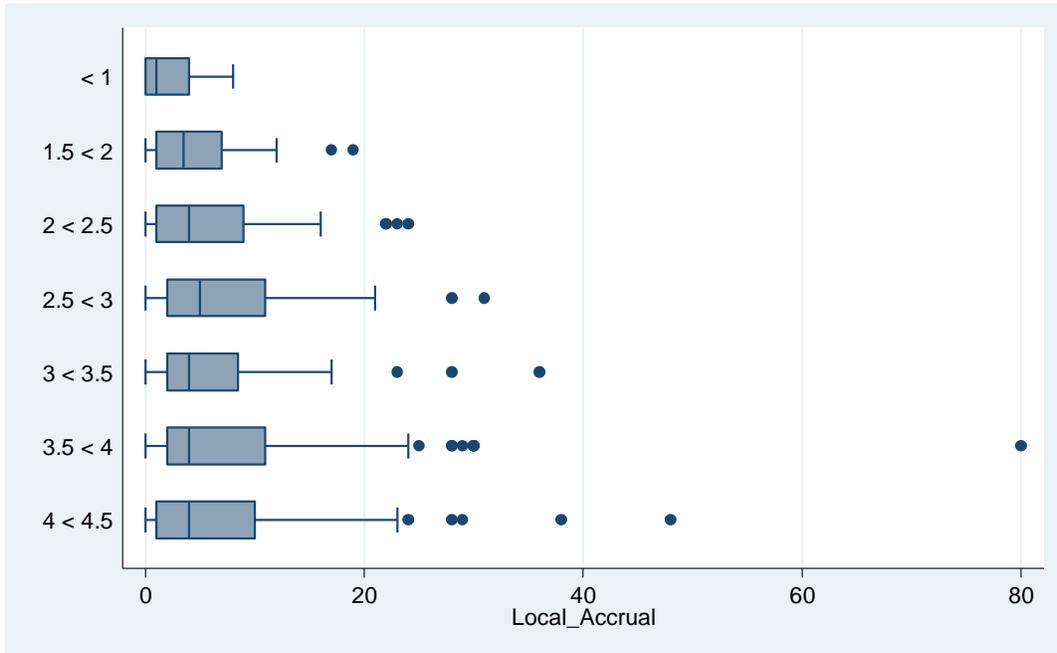


Appendix F-4g: Site 714145

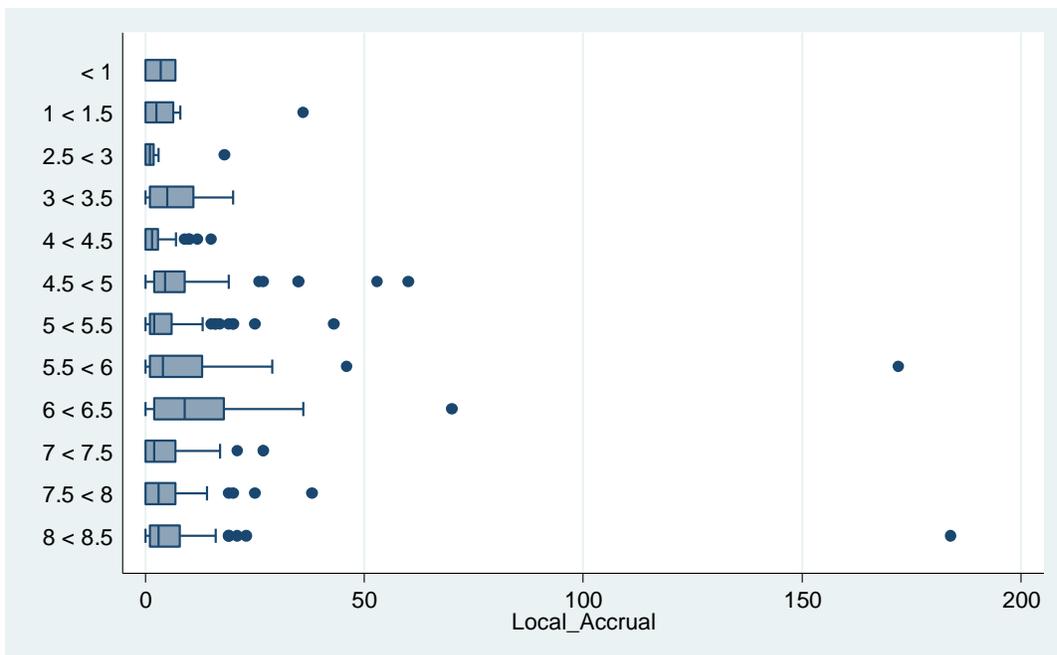


Appendix F-5: Protocols per Staff Member versus Overall Protocol Accrual

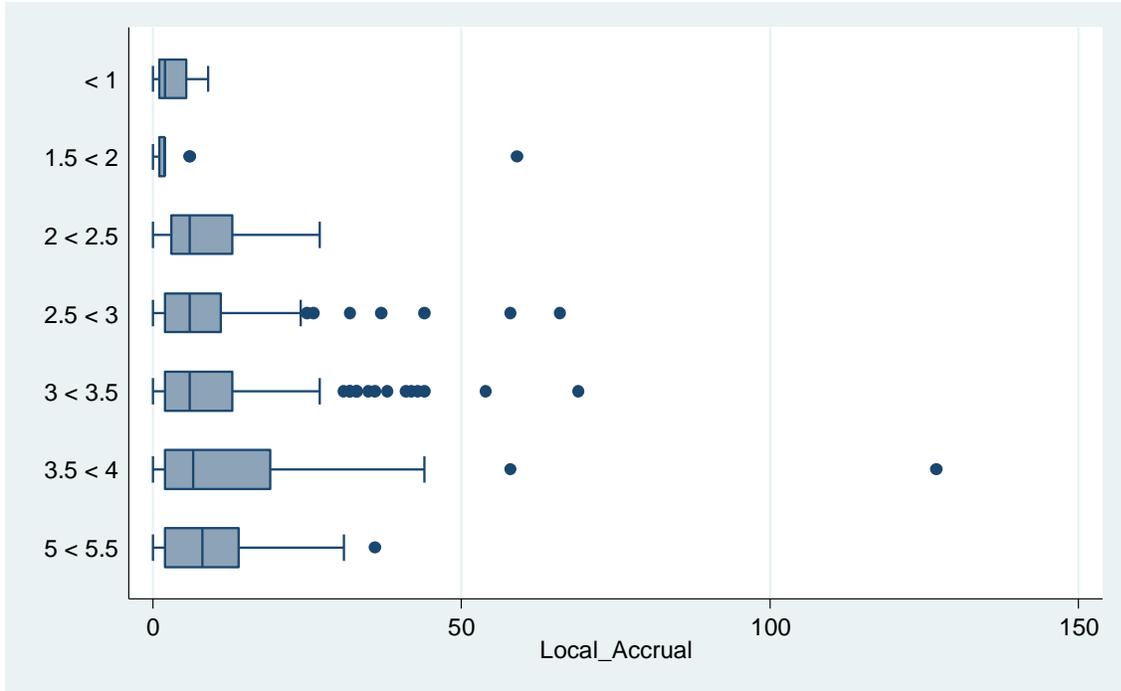
Appendix F-5a: Site 448155



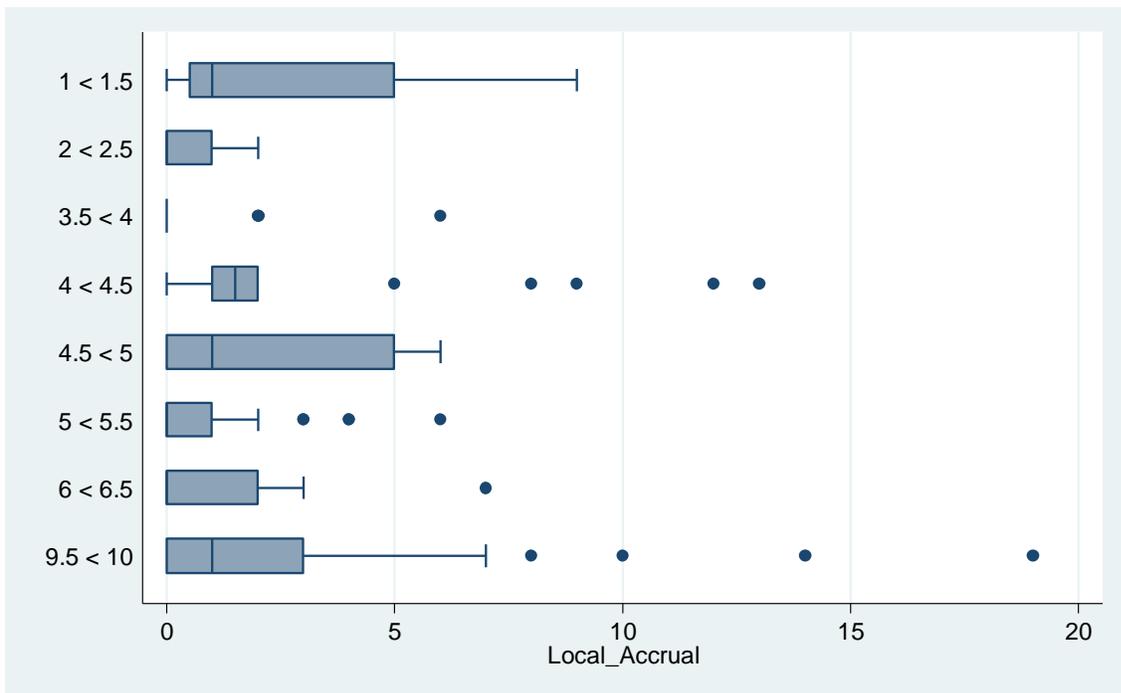
Appendix F-5b: Site 494048



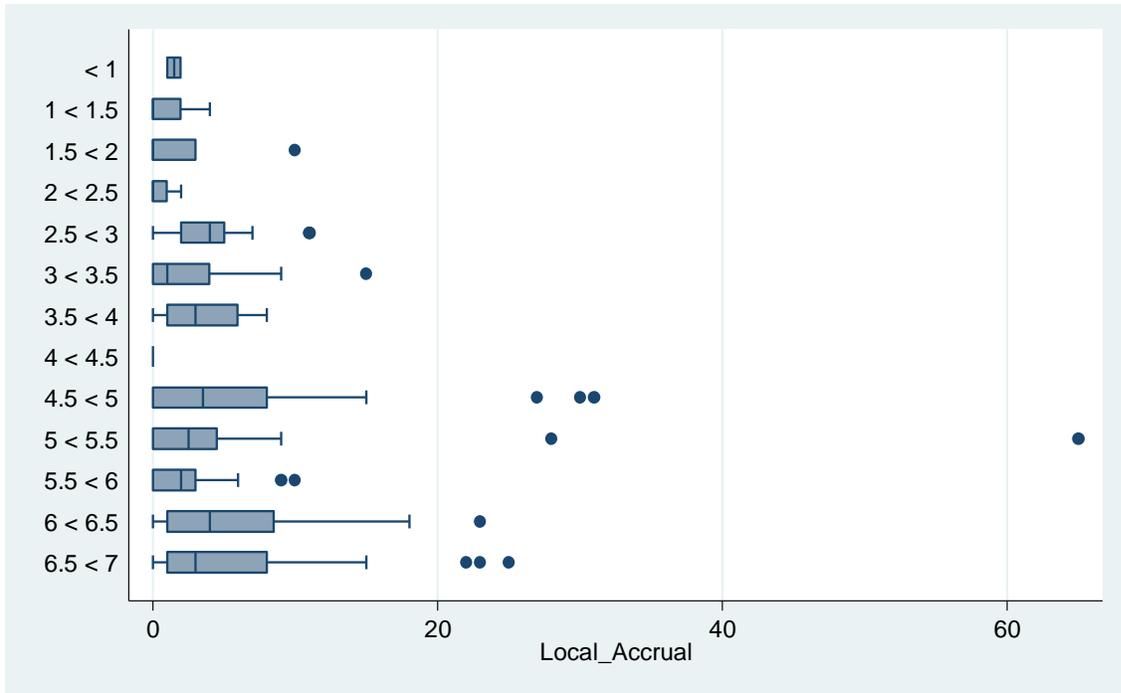
Appendix F-5c: Site 512786



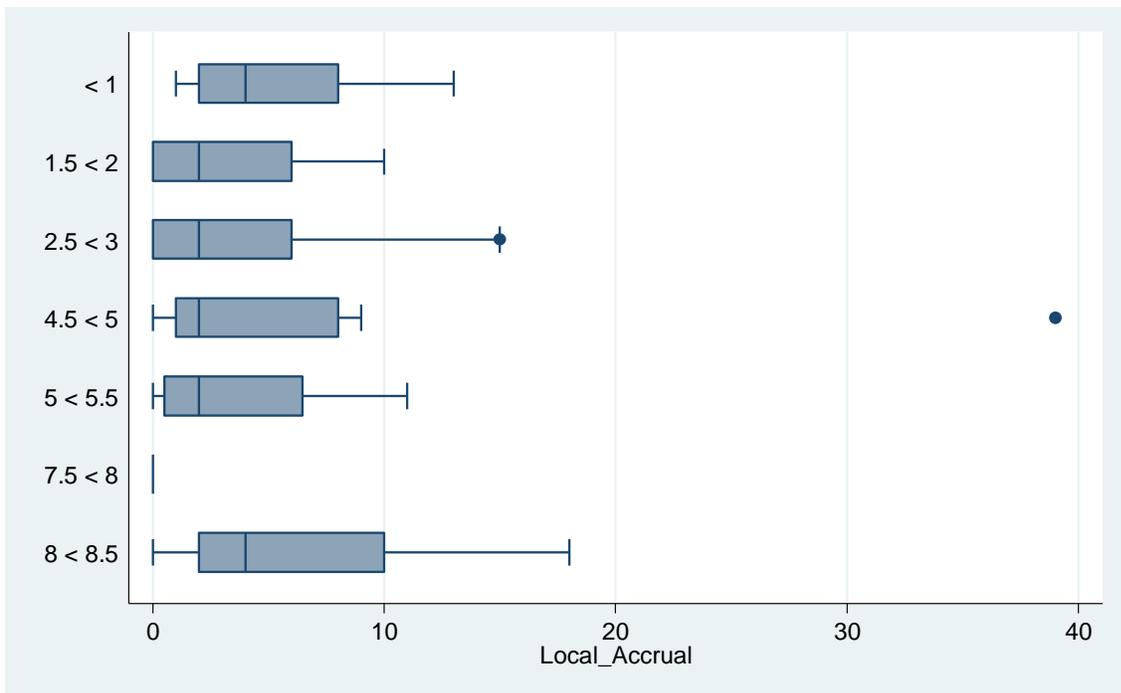
Appendix F-5d: Site 560623



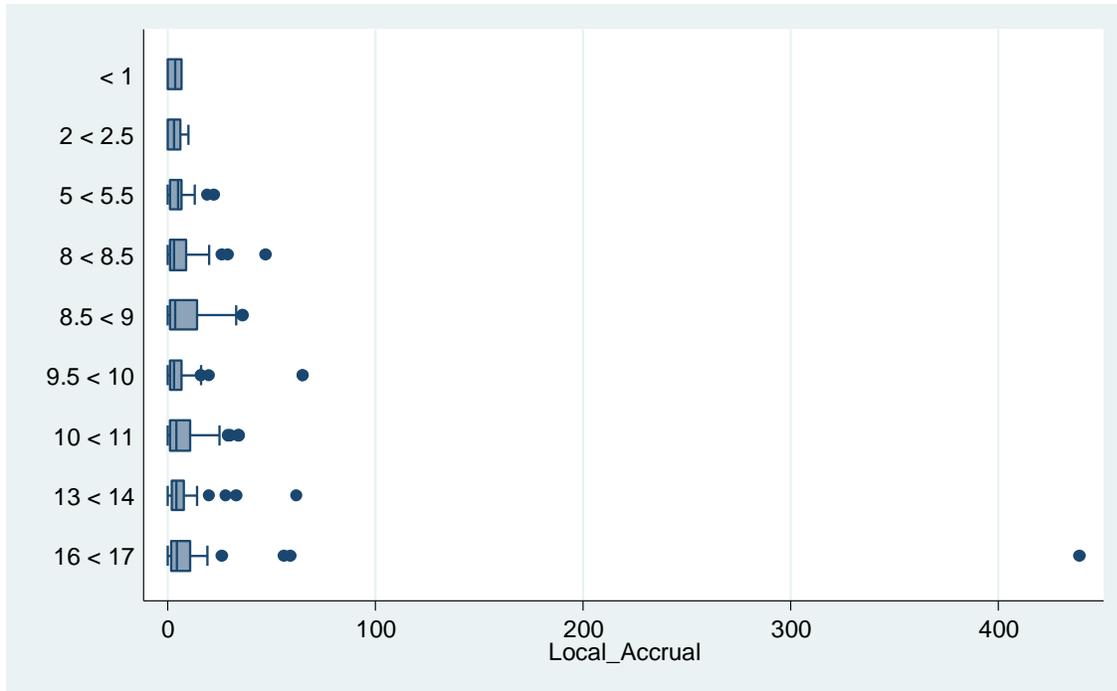
Appendix F-5e: Site 602591



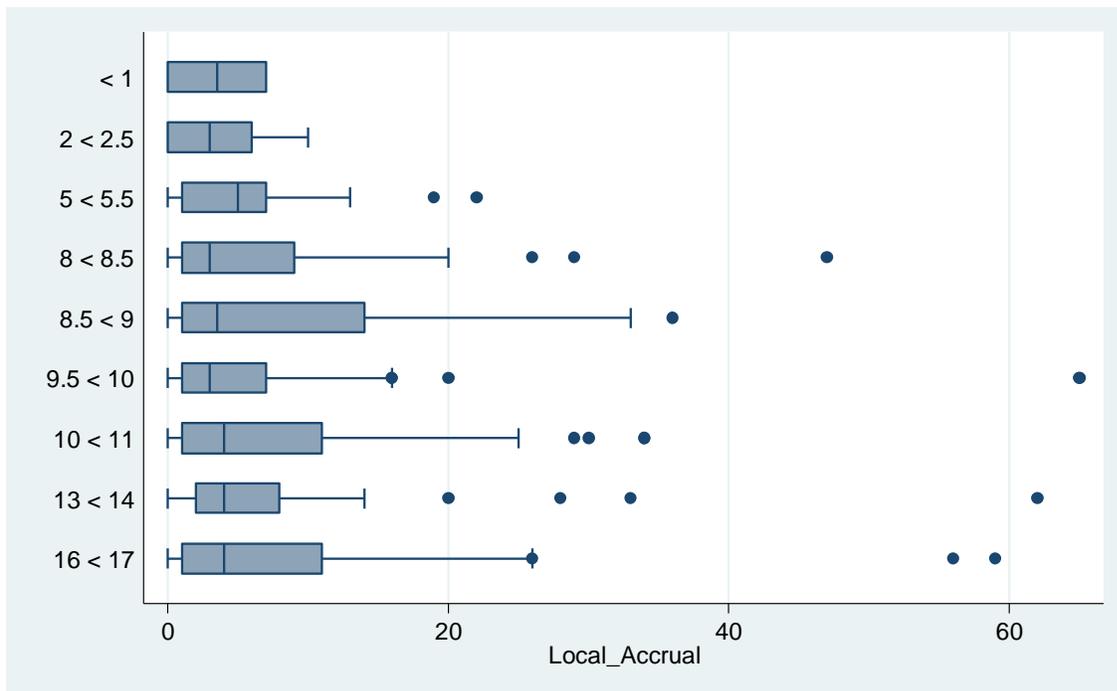
Appendix F-5f: Site 696337



Appendix F-5g: Site 714145

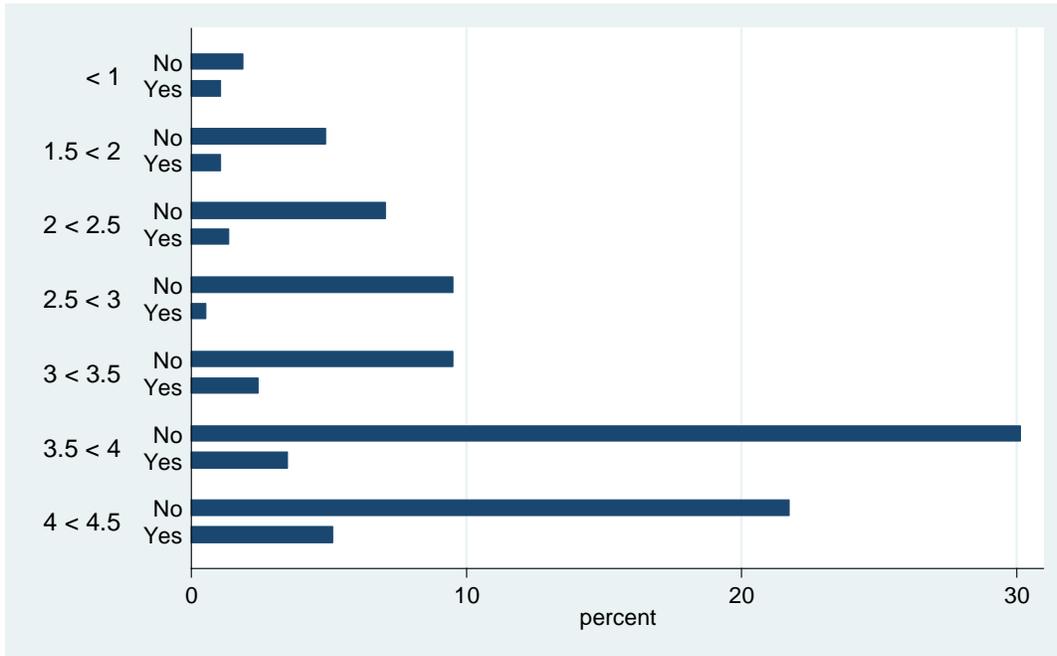


Appendix F-5h: Site 714145 – Under 100 Subjects

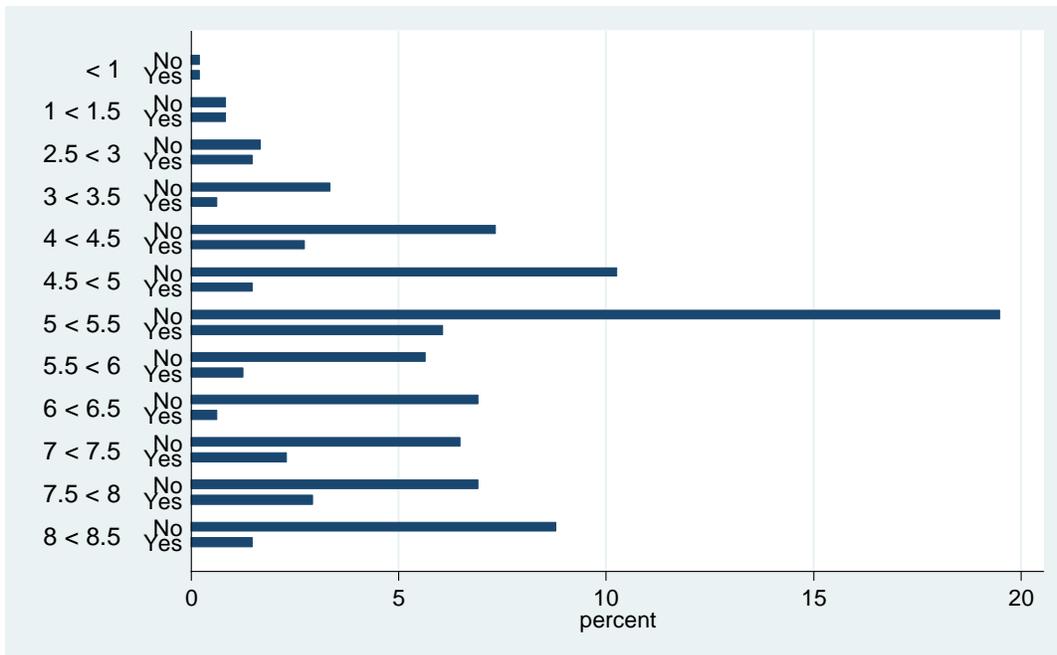


Appendix F-6: Protocols per Staff Member by Zero-Accruing Status

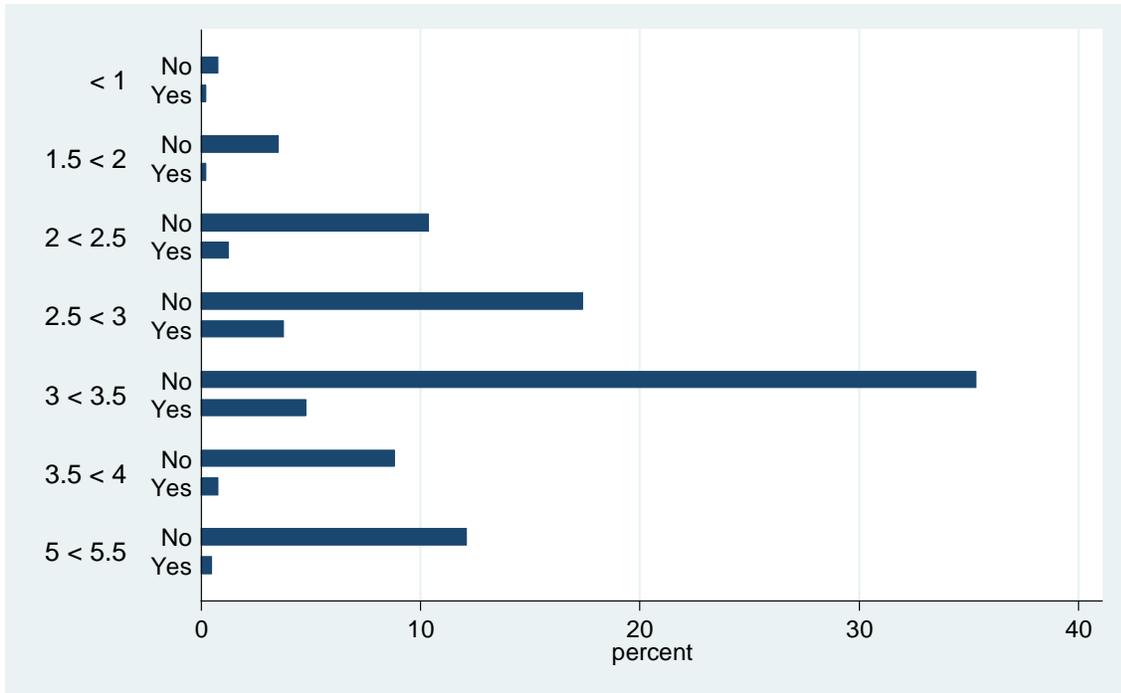
Appendix F-6a: Site 448155



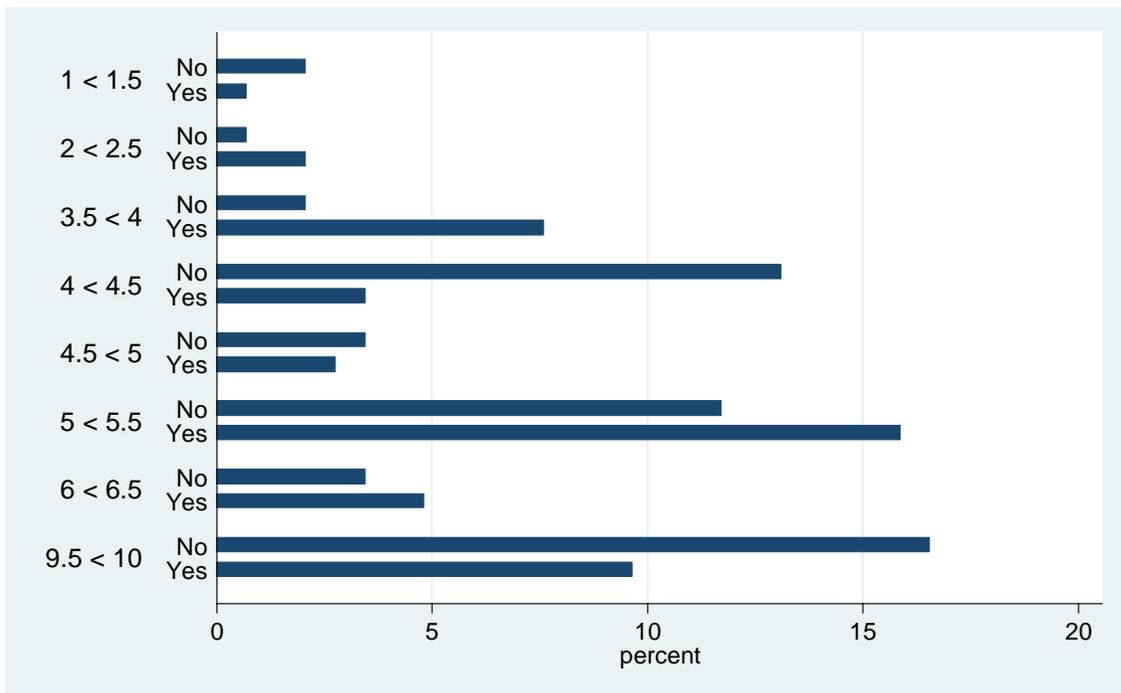
Appendix F-6b: Site 494048



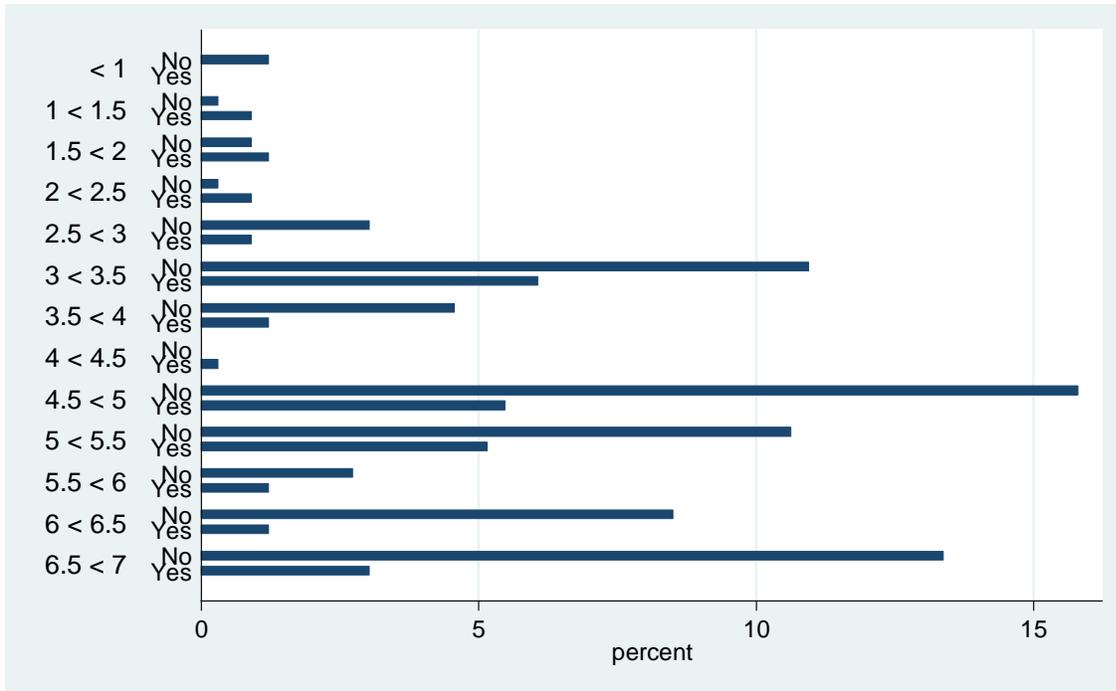
Appendix F-6c: Site 512786



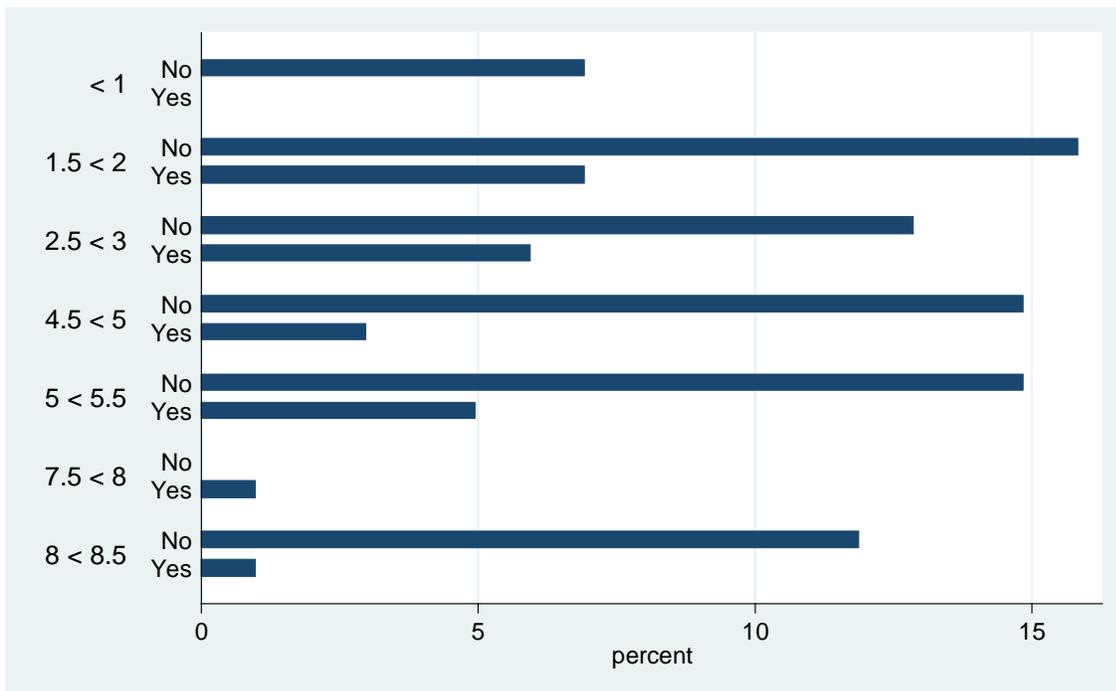
Appendix F-6d: Site 560623



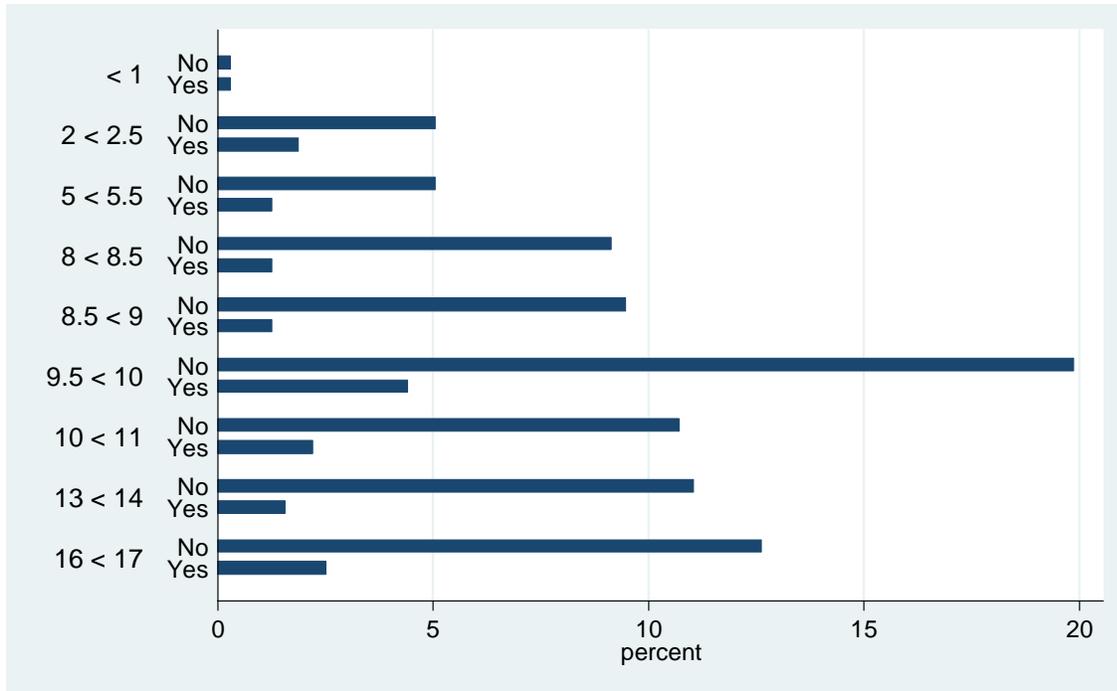
Appendix F-6e: Site 602591



Appendix F-6f: Site 696337



Appendix F-6g: Site 714145



Appendix F-7: Efficiency Outcomes by Protocol Workload

Appendix F-7a: Site 448155 – Protocols per Staff Member

# PROTOCOLS PER STAFF MEMBER	# ZERO-ACCRUING (%)	# 4+ ACCRUING (%)	AVERAGE TRIAL ACCRUAL (MEDIAN)	LOWEST MEDIAN ACTIVATION TIME
ALL STUDIES (N=368)	56 (15.2%)	199 (54.1%)	7.1 ± 8.7 (4)	167.7 ± 110.7 (155)
<1	4 (36.4%)	3 (27.3%)	2.1 ± 2.7 (1)	170.5 ± 104.2 (173)
1.5 < 2.0	4 (18.2%)	11 (50.0%)	5.0 ± 5.4 (3.5)	142.4 ± 68.2 (122)
2.0 < 2.5	5 (16.1%)	18 (58.1%)	6.8 ± 7.3 (4)	154.0 ± 104.0 (140)
2.5 < 3.0	2 (5.4%)	20 (54.1%)	7.8 ± 7.8 (5)	156.9 ± 98.8 (161)
3.0 < 3.5	9 (20.5%)	23 (52.3%)	6.2 ± 7.5 (4)	160.75 ± 119.0 (121)
3.5 < 4.0	13 (10.5%)	73 (58.9%)	8.0 ± 10.3 (4)	181.9 ± 116.2 (169)
4.0 < 4.5	19 (19.2%)	51 (51.5%)	7.1 ± 8.8 (4)	116.5 ± 114.8 (155)
NOT SPECIFIED (N=329)	44 (13.4%)	205 (62.3%)	12.5 ± 18.1 (6)	153.4 ± 158.1 (113)

Appendix F-7b: Site 448155 – Protocols per FTE

# PROTOCOLS PER FTE	# ZERO-ACCRUING (%)	# 4+ ACCRUING (%)	AVERAGE TRIAL ACCRUAL (MEDIAN)	LOWEST MEDIAN ACTIVATION TIME
ALL STUDIES (N=368)	56 (15.2%)	199 (54.1%)	7.1 ± 8.7 (4)	167.7 ± 110.7 (155)
<1	4 (36.4%)	3 (27.3%)	2.1 ± 2.6 (1)	170.5 ± 104.2 (173)
1.5 < 2.0	4 (18.2%)	11 (50.0%)	5.0 ± 5.4 (3.5)	142.4 ± 68.2 (122)
2.5 < 3.0	5 (16.1%)	18 (58.1%)	6.8 ± 7.3 (4)	154.0 ± 104.0 (140)

3.0 < 3.5	11 (13.6%)	43 (53.1%)	6.9 ± 7.6 (4)	159.0 ± 109.6 (137)
3.5 < 4.0	13 (10.5%)	73 (58.9%)	8.0 ± 10.3 (4)	181.9 ± 116.2 (169)
4.0 < 4.5	11 (23.9%)	22 (47.9%)	7.9 ± 10.8 (3)	146.9 ± 90.1 (131)
4.5 < 5.0	8 (15.1%)	29 (54.7%)	6.4 ± 6.6 (4)	183.5 ± 131.1 (175)
NOT SPECIFIED (N=329)	44 (13.4%)	205 (62.3%)	12.5 ± 18.1 (6)	153.4 ± 158.1 (113)

Appendix F-7c: Site 494048 – Protocols per Staff Member/FTE

# PROTOCOLS PER STAFF MEMBER/FTE	# ZERO- ACCRUING (%)	# 4+ ACCRUING (%)	AVERAGE TRIAL ACCRUAL (MEDIAN)	LOWEST MEDIAN ACTIVATION TIME
ALL STUDIES (N=477)	105 (22.0%)	201 (42.1%)	6.5 ± 14.1 (3)	123.9 ± 110.2 (101)
<1	1 (50.0%)	1 (50.0%)	3.5 ± 4.9 (3.5)	47 ± 21.2 (47)
1.0 < 1.5	4 (50.0%)	4 (50.0%)	6.8 ± 12.2 (2.5)	69.1 ± 37.9 (62)
2.5 < 3.0	7 (46.7%)	1 (6.7%)	2.1 ± 4.5 (1)	67.7 ± 36.3 (57)
3.0 < 3.5	3 (16.8%)	13 (68.4%)	7.1 ± 6.8 (5)	76.5 ± 41.9 (64)
4.0 < 4.5	13 (27.1%)	11 (22.9%)	2.8 ± 3.5 (1.5)	120.4 ± 61.0 (124)
4.5 < 5.0	7 (12.5%)	32 (57.1%)	8.6 ± 12.3 (4.5)	110.1 ± 122.9 (81)
5.0 < 5.5	29 (23.8%)	45 (36.9%)	4.3 ± 6.1 (2)	142.6 ± 83.5 (143.5)
5.5 < 6.0	6 (18.2%)	18 (54.6%)	13.3 ± 30.4 (4)	122.6 ± 106.2 (94)
6.0 < 6.5	3 (8.3%)	24 (66.7%)	11.4 ± 13.6 (9)	123.6 ± 145.4 (74)
7.0 < 7.5	11 (26.2%)	16 (38.1%)	5.0 ± 7.1 (2)	123.3 ± 125.1 (99.5)
7.5 < 8.0	14 (29.8%)	17 (36.2%)	5.1 ± 7.6 (3)	150 ± 184.8 (112)
8.0 < 8.5	7 (14.3%)	19 (38.8%)	9.0 ± 26.3 (3)	121.2 ± 84.6 (103)

NOT SPECIFIED (N=75)	13 (17.5%)	46 (61.3%)	12.2 ± 18.6 (5)	125.0 ± 89.7 (96)
----------------------	------------	------------	-----------------	-------------------

Appendix F-7d: Site 512786 – Protocols per Staff Member/FTE

# PROTOCOLS PER STAFF MEMBER/FTE	# ZERO-ACCRUING (%)	# 4+ ACCRUING (%)	AVERAGE TRIAL ACCRUAL (MEDIAN)	LOWEST MEDIAN ACTIVATION TIME
ALL STUDIES (N=396)	46 (11.6%)	254 (64.1%)	9.8 ± 12.7 (6)	216.0 ± 137.7 (180)
<1	1 (25.0%)	1 (25.0%)	3.3 ± 3.9 (2)	159.5 ± 38.2 (145)
1.5 < 2.0	1 (6.7%)	3 (20.0%)	5.9 ± 14.8 (2)	153.7 ± 76.3 (118)
2.0 < 2.5	5 (10.9%)	34 (73.9%)	8.3 ± 6.9 (6)	231.6 ± 167.3 (200)
2.5 < 3.0	15 (17.9%)	53 (63.1%)	9.2 ± 12.0 (6)	191.7 ± 116.2 (155)
3.0 < 3.5	19 (12.0%)	104 (65.4%)	10.0 ± 11.9 (6)	218.7 ± 116.0 (182)
3.5 < 4.0	3 (7.9%)	27 (71.1%)	14.7 ± 23.0 (6.5)	227.7 ± 156.4 (196.5)
5.0 < 5.5	2 (4.0%)	32 (64.0%)	9.1 ± 8.2 (8)	248.2 ± 191.8 (196)
NOT SPECIFIED (N=88)	0 (0.0%)	72 (81.8%)	26.9 ± 54.0 (15)	278.6 ± 338.1 (185)

Appendix F-7e: Site 560623 – Protocols per Staff Member/FTE

# PROTOCOLS PER STAFF MEMBER/FTE	# ZERO-ACCRUING (%)	# 4+ ACCRUING (%)	AVERAGE TRIAL ACCRUAL (MEDIAN)	LOWEST MEDIAN ACTIVATION TIME
ALL STUDIES (N=145)	68 (46.9%)	23 (15.9%)	1.9 ± 3.1 (1)	173.4 ± 128.3 (141)
1.0 < 1.5	1 (25.0%)	1 (25.0%)	2.8 ± 4.2 (1)	108.5 ± 15.9 (107)
2.0 < 2.5	3 (75.0%)	0 / 4 (0.0%)	0.5 ± 1.0 (0)	107 ± 101.6 (72.5)

3.5 < 4.0	11 (78.6%)	1 (7.1%)	0.7 ± 1.7 (0)	212.8 ± 178.3 (145)
4.0 < 4.5	5 (20.8%)	5 (20.8%)	2.8 ± 3.8 (1.5)	165.4 ± 99.5 (163.5)
4.5 < 5.0	4 (44.4%)	3 (33.3%)	2.0 ± 2.5 (1)	168.9 ± 92.3 (175)
5.0 < 5.5	23 (57.5%)	3 (7.5%)	0.9 ± 1.4 (0)	180.5 ± 131.6 (150)
6.0 < 6.5	7 (58.3%)	1 (8.3%)	1.3 ± 2.1 (0)	208.8 ± 101.9 (169)
9.5 < 10.0	14 (36.8%)	9 (23.7%)	2.9 ± 4.2 (1)	160.1 ± 143.2 (105)
NOT SPECIFIED (N=34)	13 (38.2%)	6 (17.7%)	2.8 ± 4.7 (1)	135.7 ± 138.3 (92)

Appendix F-7f: Site 602591 – Protocols per Staff Member

# PROTOCOLS PER STAFF MEMBER	# ZERO- ACCRUING (%)	# 4+ ACCRUING (%)	AVERAGE TRIAL ACCRUAL (MEDIAN)	LOWEST MEDIAN ACTIVATION TIME
ALL STUDIES (N=329)	91 (27.7%)	131 (39.8%)	4.1 ± 6.1 (2)	222.0 ± 135.9 (209)
<1	0 / 4 (0.0%)	0 / 4 (0.0%)	1.5 ± 0.6 (1.5)	41.3 ± 7.8 (41)
1.0 < 1.5	3 (75.0%)	1 (25.0%)	1.0 ± 2.0 (0)	142.5 ± 65.1 (151.5)
1.5 < 2.0	4 (57.1%)	1 (14.3%)	2.0 ± 3.7 (0)	95.3 ± 30.8 (105)
2.0 < 2.5	3 (75.0%)	0 / 4 (0.0%)	0.5 ± 1.0 (0)	333.3 ± 38.2 (330)
2.5 < 3.0	3 (23.1%)	7 (53.9%)	4.2 ± 3.7 (4)	298.4 ± 108.4 (305)
3.0 < 3.5	20 (35.7%)	16 (28.6%)	2.3 ± 2.8 (1)	173.9 ± 91.2 (152)
3.5 < 4.0	4 (21.1%)	8 (42.1%)	3.5 ± 2.7 (3)	209.2 ± 136.4 (141)
4.0 < 4.5	1 (100.0%)	0 / 1 (0.0%)	0 (0)	833
4.5 < 5.0	18 (25.7%)	35 (50.0%)	5.3 ± 6.7 (3.5)	220.1 ± 122.5 (222)
5.0 < 5.5	17 (32.7%)	18 (34.6%)	4.2 ± 9.6 (2.5)	258.5 ± 141.9 (243.5)

5.5 < 6.0	4 (30.8%)	3 (23.1%)	2.8 ± 3.4 (2)	404 ± 225.5 (318)
6.0 < 6.5	4 (12.5%)	18 (56.3%)	5.5 ± 5.5 (4)	227.6 ± 129.2 (222.5)
6.5 < 7.0	10 (18.5%)	24 (44.4%)	5.1 ± 5.9 (3)	194.4 ± 94.0 (179)
NOT SPECIFIED (N=66)	3 (4.6%)	44 (66.7%)	11.5 ± 10.5 (8)	210.7 ± 110.0 (178)

Appendix F-7g: Site 602591 – Protocols per FTE

# PROTOCOLS PER FTE	# ZERO- ACCRUING (%)	# 4+ ACCRUING (%)	AVERAGE TRIAL ACCRUAL (MEDIAN)	LOWEST MEDIAN ACTIVATION TIME
ALL STUDIES (N=329)	91 (27.7%)	131 (39.8%)	4.1 ± 6.1 (2)	222.0 ± 135.9 (209)
<1	0 / 4 (0.0%)	0 / 4 (0.0%)	1.5 ± 0.6 (1.5)	41.3 ± 7.8 (41)
1.0 < 1.5	3 (75.0%)	1 (25.0%)	1.0 ± 2.0 (0)	142.5 ± 65.1 (151.5)
1.5 < 2.0	4 (57.1%)	1 (14.3%)	2.0 ± 3.7 (0)	95.3 ± 30.8 (105)
2.0 < 2.5	3 (75.0%)	0 / 4 (0.0%)	0.5 ± 1.0 (0)	333.3 ± 38.2 (330)
2.5 < 3.0	3 (23.1%)	7 (53.9%)	4.2 ± 3.7 (4)	298.4 ± 108.4 (305)
3.0 < 3.5	18 (34.6%)	16 (30.8%)	2.4 ± 2.9 (1)	179.6 ± 91.9 (162)
3.5 < 4.0	2 (50.0%)	0 / 4 (0.0%)	0.5 ± 0.6 (0.5)	99.8 ± 32.2 (94.5)
4.0 < 4.5	5 (25.0%)	8 (40.0%)	3.3 ± 2.7 (3)	240.4 ± 195.6 (158.5)
5.0 < 5.5	18 (25.7%)	35 (50.0%)	5.3 ± 6.7 (3.5)	220.1 ± 122.5 (222)
5.5 < 6.0	21 (32.3%)	21 (32.3%)	4.0 ± 8.7 (2)	287.6 ± 170.4 (274)
7.0 < 7.5	8 (15.1%)	28 (52.8%)	5.0 ± 4.8 (4)	208.8 ± 113.1 (202)
7.5 < 8.0	6 (18.2%)	14 (42.4%)	5.6 ± 6.9 (2)	203.5 ± 103.4 (198)

NOT SPECIFIED (N=66) | 3 (4.6%) 44 (66.7%) 11.5 ± 10.5 (8) 210.7 ± 110.0 (178)

Appendix F-7h: Site 696337 – Protocols per Staff Member/FTE

# PROTOCOLS PER STAFF MEMBER/FTE	# ZERO-ACCRUING (%)	# 4+ ACCRUING (%)	AVERAGE TRIAL ACCRUAL (MEDIAN)	LOWEST MEDIAN ACTIVATION TIME
ALL STUDIES (N=101)	23 (22.8%)	43 (42.6%)	4.3 ± 5.4 (2)	155.6 ± 114.2 (137)
<1	0 / 7 (0.0%)	4 (57.1%)	5.3 ± 4.2 (4)	139 ± 118.6 (90)
1.5 < 2.0	7 (30.4%)	10 (43.5%)	3.4 ± 3.6 (2)	204.4 ± 81.2 (215)
2.5 < 3.0	6 (31.6%)	7 (36.8%)	3.5 ± 4.2 (2)	217.6 ± 113.1 (173)
4.5 < 5.0	3 (16.7%)	8 (44.4%)	5.6 ± 9.0 (2)	88 ± 88.1 (57)
5.0 < 5.5	5 (25.0%)	7 (35.0%)	3.4 ± 3.6 (2)	189.4 ± 113.7 (152.5)
7.5 < 8.0	1 (100.0%)	0 / 1 (0.0%)	0	Not provided
8.0 < 8.5	1 (7.8%)	7 (53.9%)	6.5 ± 5.6 (4)	29.1 ± 28.2 (20)
NOT SPECIFIED (N=81)	15 (18.5%)	34 (42.0%)	4.9 ± 6.0 (3)	Not provided

Appendix F-7i: Site 714145 – Protocols per Staff Member/FTE

# PROTOCOLS PER STAFF MEMBER/FTE	# ZERO-ACCRUING (%)	# 4+ ACCRUING (%)	AVERAGE TRIAL ACCRUAL (MEDIAN)	LOWEST MEDIAN ACTIVATION TIME
ALL STUDIES (N=317)	53 (16.7%)	161 (50.8%)	8.2 ± 26.3 (4)	180.0 ± 138.3 (149)
<1	1 (50.0%)	1 (50.0%)	3.5 ± 4.9 (3.5)	59.5 ± 13.4 (59.5)
2.0 < 2.5	6 (27.3%)	10 (45.5%)	3.2 ± 3.0 (3)	187.0 ± 81.1 (180)

5.0 < 5.5	4 (20.0%)	14 (70.0%)	5.9 ± 6.1 (5)	159.5 ± 86.4 (160.5)
8.0 < 8.5	4 (12.1%)	13 (39.4%)	7.1 ± 10.2 (3)	204.1 ± 113.9 (196)
8.5 < 9.0	4 (11.8%)	17 (50.0%)	8.6 ± 10.9 (3.5)	166.8 ± 130.2 (149)
9.5 < 10.0	14 (18.2%)	33 (42.9%)	5.7 ± 10.7 (3)	221.4 ± 179.6 (181)
10 < 11	7 (17.1%)	23 (56.1%)	7.8 ± 9.8 (4)	150.1 ± 132.1 (97)
13 < 14	5 (12.5%)	24 (60.0%)	7.5 ± 11.3 (4)	155.9 ± 115.9 (135.5)
16 < 17	8 (16.7%)	26 (54.2%)	17.0 ± 63.3 (4.5)	162.5 ± 134.0 (113)
NOT SPECIFIED (N=212)	27 (12.7%)	136 (64.2%)	13.8 ± 26.9 (6)	192.1 ± 172.0 (134)

Appendix F-8 Site-Specific Full Models Assessing the Impact of Protocol Workload on Clinical Trial Accrual

Appendix F-8a: Site 494048 – Protocols per Staff Member and Protocols per FTE

```

Negative binomial regression      Number of obs   =      477
                                 LR chi2(48)      =     316.91
Dispersion      = mean          Prob > chi2     =     0.0000
Log likelihood = -1205.7418     Pseudo R2      =     0.1162
    
```

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

prot_per_fte_cat						
1 < 1.5	.9536286	.9047947	1.05	0.292	-.8197365 2.726994	
2.5 < 3	.1885061	.8916472	0.21	0.833	-1.55909 1.936102	
3 < 3.5	1.161936	.8645961	1.34	0.179	-.532641 2.856514	
4 < 4.5	.4692361	.8363413	0.56	0.575	-1.169963 2.108435	
4.5 < 5	1.205612	.8321165	1.45	0.147	-.4253061 2.836531	
5 < 5.5	.746071	.8212074	0.91	0.364	-.863466 2.355608	
5.5 < 6	1.225865	.8449262	1.45	0.147	-.4301596 2.88189	
6 < 6.5	1.147494	.8478038	1.35	0.176	-.5141712 2.809159	
7 < 7.5	.7132664	.8341279	0.86	0.392	-.9215943 2.348127	
7.5 < 8	.761909	.8366663	0.91	0.362	-.8779268 2.401745	
8 < 8.5	.847996	.8407589	1.01	0.313	-.7998611 2.495853	
phase						
Phase 1	.6969194	1.009155	0.69	0.490	-1.280988 2.674827	
Phase 1/2	.8622798	1.017598	0.85	0.397	-1.132175 2.856735	
Phase 2	.4835818	1.007614	0.48	0.631	-1.491306 2.458469	
Phase 2/3	.4106648	1.127738	0.36	0.716	-1.799661 2.62099	
Phase 3	.1409548	1.024	0.14	0.891	-1.866049 2.147958	
Phase 4	.9062787	1.11965	0.81	0.418	-1.288194 3.100751	
None	.68271	1.047097	0.65	0.514	-1.369562 2.734982	
sponsor_type						
Institutitonal	.1049192	.3497298	0.30	0.764	-.5805386 .790377	
Industry	-.6969732	.3441344	-2.03	0.043	-1.371464 -.0224821	
National Group	-.9212774	.3684051	-2.50	0.012	-1.643338 -.1992166	
peds						
Yes	-.243695	.1884198	-1.29	0.196	-.612991 .1256009	
placebo						
Yes	-.3183175	.1800068	-1.77	0.077	-.6711244 .0344894	

natenroll_trans		.6316371	.0706836	8.94	0.000	.4930998	.7701744
natsitescat							
10-49 sites		-.5908104	.1602119	-3.69	0.000	-.9048199	-.2768009
50-199 sites		-.9838104	.2206099	-4.46	0.000	-1.416198	-.551423
200+ sites		-1.014545	.3114802	-3.26	0.001	-1.625035	-.4040546
Unknown		-.7522424	.4310281	-1.75	0.081	-1.597042	.0925573
totalmo_trans		.1003965	.1149235	0.87	0.382	-.1248496	.3256425
moaccrdone_trans		-.219776	.0440746	-4.99	0.000	-.3061607	-.1333914
primary_purpose							
Supportive Care		.2807803	.3120005	0.90	0.368	-.3307294	.89229
disease_team							
1		-1.584976	.6786247	-2.34	0.020	-2.915056	-.2548959
2		-1.329791	.6524956	-2.04	0.042	-2.608659	-.0509229
4		-2.284423	.7042664	-3.24	0.001	-3.66476	-.9040859
5		-1.019157	.6755843	-1.51	0.131	-2.343278	.3049639
6		-1.163447	1.248562	-0.93	0.351	-3.610585	1.28369
7		-1.209482	.6721665	-1.80	0.072	-2.526904	.1079402
8		-.6662743	.7439731	-0.90	0.370	-2.124435	.7918862
9		-.619418	.6984691	-0.89	0.375	-1.988392	.7495563
10		-2.084177	.6772164	-3.08	0.002	-3.411496	-.756857
11		-1.206393	.7712589	-1.56	0.118	-2.718033	.3052468
12		-1.505061	.6752219	-2.23	0.026	-2.828472	-.1816504
13		-2.816497	.7891796	-3.57	0.000	-4.36326	-1.269733
14		-1.601681	.6694041	-2.39	0.017	-2.913689	-.2896727
15		-1.314411	.6994641	-1.88	0.060	-2.685335	.0565134
16		-2.041894	.849228	-2.40	0.016	-3.70635	-.3774376
18		-1.195755	.7639917	-1.57	0.118	-2.693151	.3016412
19		-1.313435	.6657707	-1.97	0.049	-2.618322	-.0085485
_cons		-.0028507	1.502216	-0.00	0.998	-2.947139	2.941438

/lnalpha		-.2395553	.093366			-.4225494	-.0565612

alpha		.7869777	.073477			.6553739	.9450086

LR test of alpha=0: chibar2(01) = 1041.95				Prob >= chibar2 = 0.000			

Measures of Fit for nbreg of local_accrual

Log-Lik Intercept Only:	-1364.198	Log-Lik Full Model:	-1205.742
D(419):	2411.484	LR(48):	316.913
		Prob > LR:	0.000
McFadden's R2:	0.116	McFadden's Adj R2:	0.074
Maximum Likelihood R2:	0.485	Cragg & Uhler's R2:	0.487
AIC:	5.299	AIC*n:	2527.484
BIC:	-172.706	BIC':	-20.872

Appendix F-8b: Site 560623 – Protocols per Staff Member and Protocols per FTE

Negative binomial regression	Number of obs	=	145
	LR chi2(19)	=	62.59
Dispersion = mean	Prob > chi2	=	0.0000
Log likelihood = -228.22908	Pseudo R2	=	0.1206

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

prot_per_fte_cat						
2 < 2.5	.1611895	1.114398	0.14	0.885	-2.022991	2.34537
3.5 < 4	-.5131816	.8003023	-0.64	0.521	-2.081745	1.055382
4 < 4.5	.5712318	.6940362	0.82	0.410	-.789054	1.931518
4.5 < 5	1.265078	.7746927	1.63	0.102	-.2532914	2.783448
5 < 5.5	-.1219529	.6839937	-0.18	0.858	-1.462556	1.21865
6 < 6.5	.3269431	.7495145	0.44	0.663	-1.142078	1.795964
9.5 < 10	.4653249	.6576407	0.71	0.479	-.8236271	1.754277

priend						
Efficacy	.4918539	.8657514	0.57	0.570	-1.204988	2.188695
Other	.3438816	1.711662	0.20	0.841	-3.010915	3.698678

moaccrdone_trans	-.3165554	.0884519	-3.58	0.000	-.489918	-.1431929

disease_team						
1	.7074968	.4604515	1.54	0.124	-.1949715	1.609965
2	.9733188	.6365711	1.53	0.126	-.2743377	2.220975
3	.9106945	.4476156	2.03	0.042	.033384	1.788005
4	.1287087	.5357039	0.24	0.810	-.9212517	1.178669
5	1.751801	.5879766	2.98	0.003	.5993879	2.904214
6	-.7841876	.6398981	-1.23	0.220	-2.038365	.4699896
7	.3490491	.4755257	0.73	0.463	-.5829642	1.281062
9	1.927234	.8907846	2.16	0.031	.1813286	3.67314

10		-16.8154	4640.611	-0.00	0.997	-9112.246	9078.615
_cons		.0032611	1.127977	0.00	0.998	-2.207534	2.214056
/lnalpha		-.1765163	.2615985			-.6892398	.3362073
alpha		.8381851	.2192679			.5019575	1.399629

LR test of alpha=0: chibar2(01) = 61.28 Prob >= chibar2 = 0.000

Measures of Fit for nbreg of local_accrual

Log-Lik Intercept Only:	-259.523	Log-Lik Full Model:	-228.229
D(121):	456.458	LR(19):	62.588
		Prob > LR:	0.000
McFadden's R2:	0.121	McFadden's Adj R2:	0.028
Maximum Likelihood R2:	0.351	Cragg & Uhler's R2:	0.361
AIC:	3.479	AIC*n:	504.458
BIC:	-145.727	BIC':	31.970

Appendix F-8c: Site 602591 – Protocols per Staff Member

Negative binomial regression	Number of obs	=	329
	LR chi2(39)	=	224.75
Dispersion = mean	Prob > chi2	=	0.0000
Log likelihood = -710.71761	Pseudo R2	=	0.1365

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
prot_per_staff_cat					
1 < 1.5	.4355722	1.041298	0.42	0.676	-1.605334 2.476478
1.5 < 2	21.10869	13744.49	0.00	0.999	-26917.6 26959.82
2 < 2.5	20.68637	13744.49	0.00	0.999	-26918.02 26959.4
2.5 < 3	22.38413	13744.49	0.00	0.999	-26916.33 26961.09
3 < 3.5	22.05137	13744.49	0.00	0.999	-26916.66 26960.76
3.5 < 4	22.42418	13744.49	0.00	0.999	-26916.29 26961.13
4 < 4.5	2.218529	30198.98	0.00	1.000	-59186.7 59191.13
4.5 < 5	22.45964	13744.49	0.00	0.999	-26916.25 26961.17
5 < 5.5	22.34716	13744.49	0.00	0.999	-26916.36 26961.06
5.5 < 6	22.29859	13744.49	0.00	0.999	-26916.41 26961.01
6 < 6.5	22.50158	13744.49	0.00	0.999	-26916.21 26961.21
6.5 < 7	22.57464	13744.49	0.00	0.999	-26916.14 26961.29
phase					

Phase 1/2	-.3529142	.2485025	-1.42	0.156	-.8399701	.1341418
Phase 2	-.2691566	.1957031	-1.38	0.169	-.6527277	.1144145
Phase 2/3	-1.197993	.7460854	-1.61	0.108	-2.660294	.2643071
Phase 3	-.8102761	.2881272	-2.81	0.005	-1.374995	-.2455572
Phase 4	.2877521	.8950778	0.32	0.748	-1.466568	2.042072
None	-.0668291	.5211577	-0.13	0.898	-1.088279	.9546212
irb						
Local	-21.41829	13744.49	-0.00	0.999	-26960.13	26917.29
sponsor_type						
Institutional	-.6098623	.335234	-1.82	0.069	-1.266909	.0471843
Industry	-1.378635	.3115856	-4.42	0.000	-1.989332	-.7679385
National Group	-1.421024	.3406836	-4.17	0.000	-2.088752	-.7532969
peds						
Yes	-.0932505	.343935	-0.27	0.786	-.7673507	.5808496
natenroll_trans	.8917764	.2337858	3.81	0.000	.4335647	1.349988
natsitescat						
10-49 sites	-.2716364	.1848727	-1.47	0.142	-.6339802	.0907075
50-199 sites	-.6435987	.2376334	-2.71	0.007	-1.109352	-.1778458
200+ sites	-.6777797	.3070073	-2.21	0.027	-1.279503	-.0760565
Unknown	-.4170083	.382398	-1.09	0.275	-1.166495	.332478
totalmo_trans	.4282352	.3348587	1.28	0.201	-.2280759	1.084546
moaccrdone_trans	-.3696156	.0451508	-8.19	0.000	-.4581095	-.2811216
precision						
Yes	-.4086533	.3302405	-1.24	0.216	-1.055913	.2386062
Conditional	-.2565285	.4009181	-0.64	0.522	-1.042314	.5292566
disease_team						
1	1.329202	.6683879	1.99	0.047	.019186	2.639218
2	2.101062	.9982623	2.10	0.035	.1445035	4.05762
3	1.285066	.6445976	1.99	0.046	.0216783	2.548454
4	.5593174	.6676817	0.84	0.402	-.7493148	1.867949
5	.7981471	.6691429	1.19	0.233	-.5133489	2.109643
6	.197566	.7587315	0.26	0.795	-1.28952	1.684653
7	1.224875	.6495146	1.89	0.059	-.0481505	2.4979
8	.8259443	.6662942	1.24	0.215	-.4799682	2.131857
9	.8269919	.653236	1.27	0.206	-.453327	2.107311
10	1.126976	.6811867	1.65	0.098	-.2081251	2.462078
11	1.451936	.6587666	2.20	0.028	.1607774	2.743095
12	1.534025	.6722706	2.28	0.022	.2163988	2.851651

13		.1563275	.6885226	0.23	0.820	-1.193152	1.505807
natenroll_trans		0	(omitted)				
totalmo_trans		0	(omitted)				
c.natenroll_trans#c.totalmo_trans		-.073994	.0662234	-1.12	0.264	-.2037896	.0558015
_cons		-2.191619	1.622203	-1.35	0.177	-5.371078	.9878394

/lnalpha		-.7876237	.15677			-1.094887	-.4803601

alpha		.4549246	.0713185			.3345773	.6185606

LR test of alpha=0: chibar2(01) = 203.44 Prob >= chibar2 = 0.000

Measures of Fit for nbreg of local_accrual

Log-Lik Intercept Only:	-823.095	Log-Lik Full Model:	-710.718
D(271):	1421.435	LR(39):	224.754
		Prob > LR:	0.000
McFadden's R2:	0.137	McFadden's Adj R2:	0.066
Maximum Likelihood R2:	0.495	Cragg & Uhler's R2:	0.498
AIC:	4.673	AIC*n:	1537.435
BIC:	-149.296	BIC':	1.292

Appendix F-8d: Site 602591 – Protocols per FTE

Negative binomial regression	Number of obs	=	329
	LR chi2(34)	=	216.38
Dispersion = mean	Prob > chi2	=	0.0000
Log likelihood = -714.90273	Pseudo R2	=	0.1314

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
prot_per_fte_cat					
1 < 1.5	.5005978	1.04919	0.48	0.633	-1.555777 2.556973
1.5 < 2	20.0726	7268.626	0.00	0.998	-14226.17 14266.32
2 < 2.5	19.50053	7268.626	0.00	0.998	-14226.75 14265.75
2.5 < 3	21.35525	7268.626	0.00	0.998	-14224.89 14267.6
3 < 3.5	21.12558	7268.626	0.00	0.998	-14225.12 14267.37
3.5 < 4	20.08756	7268.626	0.00	0.998	-14226.16 14266.33
4 < 4.5	21.34515	7268.626	0.00	0.998	-14224.9 14267.59
5 < 5.5	21.37976	7268.626	0.00	0.998	-14224.87 14267.63
5.5 < 6	21.26462	7268.626	0.00	0.998	-14224.98 14267.51

7 < 7.5		21.57395	7268.626	0.00	0.998	-14224.67	14267.82
7.5 < 8		21.47611	7268.626	0.00	0.998	-14224.77	14267.72
irb							
Local		-20.17073	7268.626	-0.00	0.998	-14266.42	14226.08
sponsor_type							
Instititutional		-.6614184	.3369661	-1.96	0.050	-1.32186	-.000977
Industry		-1.408746	.3078835	-4.58	0.000	-2.012186	-.805305
National Group		-1.382104	.3366911	-4.10	0.000	-2.042007	-.722202
random							
Yes		-.2146957	.1459713	-1.47	0.141	-.5007941	.0714028
natenroll_trans		.9431243	.2335713	4.04	0.000	.4853328	1.400916
natsitescat							
10-49 sites		-.2655085	.172605	-1.54	0.124	-.6038081	.0727912
50-199 sites		-.7569967	.2192378	-3.45	0.001	-1.186695	-.3272985
200+ sites		-.8591092	.2887954	-2.97	0.003	-1.425138	-.2930807
Unknown		-.4352206	.3796118	-1.15	0.252	-1.179246	.3088049
totalmo_trans		.5803362	.3307142	1.75	0.079	-.0678518	1.228524
moaccrdone_trans		-.3703547	.0448434	-8.26	0.000	-.4582461	-.2824633
precision							
Yes		-.3471918	.3223276	-1.08	0.281	-.9789423	.2845587
Conditional		-.2843027	.3997992	-0.71	0.477	-1.067895	.4992894
disease_team							
1		1.519903	.6645287	2.29	0.022	.217451	2.822356
2		2.370912	1.002169	2.37	0.018	.4066965	4.335126
3		1.425909	.6436999	2.22	0.027	.16428	2.687537
4		.7479728	.6612291	1.13	0.258	-.5480125	2.043958
5		.8937788	.67104	1.33	0.183	-.4214355	2.208993
6		.427352	.7779843	0.55	0.583	-1.097469	1.952173
7		1.332576	.6456428	2.06	0.039	.0671398	2.598013
8		.9906838	.6635643	1.49	0.135	-.3098783	2.291246
9		.9684915	.6530935	1.48	0.138	-.3115483	2.248531
10		1.28717	.6844995	1.88	0.060	-.0544248	2.628764
11		1.573007	.6572036	2.39	0.017	.2849117	2.861103
12		1.541914	.6660767	2.31	0.021	.2364281	2.847401
13		.250114	.6851317	0.37	0.715	-1.092719	1.592947
natenroll_trans		0	(omitted)				
totalmo_trans		0	(omitted)				

c.natenroll_trans#c.totalmo_trans		-1.077013	.0650591	-1.66	0.098	-.2352148	.0198123
_cons		-2.936024	1.598228	-1.84	0.066	-6.068494	.1964461

/lnalpha		-.7370369	.1540407			-1.038951	-.4351226

alpha		.4785298	.0737131			.3538256	.6471853

LR test of alpha=0: chibar2(01) = 217.25 Prob >= chibar2 = 0.000

Measures of Fit for nbreg of local_accrual

Log-Lik Intercept Only:	-823.095	Log-Lik Full Model:	-714.903
D(279):	1429.805	LR(34):	216.384
		Prob > LR:	0.000
McFadden's R2:	0.131	McFadden's Adj R2:	0.071
Maximum Likelihood R2:	0.482	Cragg & Uhler's R2:	0.485
AIC:	4.650	AIC*n:	1529.805
BIC:	-187.295	BIC':	-19.318

Appendix F-8e: Site 714145 – Protocols per Staff Member and Protocols per FTE

Negative binomial regression	Number of obs	=	317
	LR chi2(46)	=	333.39
Dispersion = mean	Prob > chi2	=	0.0000
Log likelihood = -805.88619	Pseudo R2	=	0.1714

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]

prot_per_staff_cat					
2 < 2.5	.521127	.7067332	0.74	0.461	-.8640446 1.906299
5 < 5.5	.5950596	.7069991	0.84	0.400	-.7906332 1.980752
8 < 8.5	.7942624	.6982877	1.14	0.255	-.5743564 2.162881
8.5 < 9	.7452319	.6944325	1.07	0.283	-.6158308 2.106295
9.5 < 10	.4352056	.6881951	0.63	0.527	-.9136319 1.784043
10 < 11	.7484266	.6912226	1.08	0.279	-.6063448 2.103198
13 < 14	.9943071	.6972494	1.43	0.154	-.3722766 2.360891
16 < 17	.4314038	.6943759	0.62	0.534	-.9295479 1.792356

phase					
Phase 1/2	.0275106	.2465958	0.11	0.911	-.4558082 .5108294
Phase 2	-.1094595	.1843173	-0.59	0.553	-.4707147 .2517957

Phase 2/3	-.5037951	.3954306	-1.27	0.203	-1.278825	.2712347
Phase 3	-.3459914	.247524	-1.40	0.162	-.8311294	.1391466
Phase 4	.1675034	.5313869	0.32	0.753	-.8739957	1.209002
None	-.4898211	.3237454	-1.51	0.130	-1.12435	.1447082
sponsor_type						
Instituitonal	.2645321	.3101779	0.85	0.394	-.3434053	.8724696
Industry	-.6670431	.3136664	-2.13	0.033	-1.281818	-.0522683
National Group	-.7002585	.4284885	-1.63	0.102	-1.540081	.1395636
peds						
Yes	.6301915	.2563876	2.46	0.014	.127681	1.132702
natenroll_trans	.8578149	.0731898	11.72	0.000	.7143654	1.001264
natsitescat						
10-49 sites	-.8126378	.1990285	-4.08	0.000	-1.202726	-.4225491
50-199 sites	-1.186121	.2305922	-5.14	0.000	-1.638074	-.7341688
200+ sites	-1.988735	.3029501	-6.56	0.000	-2.582507	-1.394964
Unknown	-1.309369	.5139281	-2.55	0.011	-2.31665	-.3020887
moaccrdone_trans	-.0173025	.1391234	-0.12	0.901	-.2899793	.2553743
primary_purpose						
Supportive Care	-.0110874	.2696311	-0.04	0.967	-.5395545	.5173798
disease_team						
1	-.05214	.3869057	-0.13	0.893	-.8104613	.7061813
2	.2440747	.3801695	0.64	0.521	-.5010438	.9891931
3	-1.275452	.4517089	-2.82	0.005	-2.160785	-.390119
4	-.0380662	.3441962	-0.11	0.912	-.7126784	.6365459
5	.0417976	.3756913	0.11	0.911	-.6945438	.7781391
6	.076909	.3941586	0.20	0.845	-.6956278	.8494457
7	-.2354716	.4404333	-0.53	0.593	-1.098705	.6277619
8	-.082903	.3362516	-0.25	0.805	-.741944	.5761381
9	-.662904	.3454242	-1.92	0.055	-1.339923	.014115
10	-.4169342	.4492494	-0.93	0.353	-1.297447	.4635784
11	-.0673223	.2933466	-0.23	0.818	-.642271	.5076263
12	-.6969512	.3184561	-2.19	0.029	-1.321114	-.0727888
13	-.3808062	.3732496	-1.02	0.308	-1.112362	.3507495
14	-1.399318	.4098996	-3.41	0.001	-2.202707	-.5959299
15	1.185236	.5745906	2.06	0.039	.0590591	2.311413
16	.1253366	.3900978	0.32	0.748	-.6392409	.8899142
17	.9611767	.4973851	1.93	0.053	-.0136802	1.936034
18	-.4859613	.3776767	-1.29	0.198	-1.226194	.2542714
19	-.5325126	.6093418	-0.87	0.382	-1.7268	.6617753

20		-.2181204	.36162	-0.60	0.546	-.9268827	.4906418
natenroll_trans		0	(omitted)				
moaccrdone_trans		0	(omitted)				
c.natenroll_trans#c.moaccrdone_trans		-.0715782	.0265936	-2.69	0.007	-.1237006	-.0194558
_cons		-.4107221	.8419946	-0.49	0.626	-2.061001	1.239557

/lnalpha		-.7631552	.1274164			-1.012887	-.5134236

alpha		.4661932	.0594007			.3631691	.5984432

LR test of alpha=0: $\chi^2(01) = 433.57$ Prob $\geq \chi^2 = 0.000$

Measures of Fit for nbreg of local_accrual

Log-Lik Intercept Only:	-972.580	Log-Lik Full Model:	-805.886
D(260):	1611.772	LR(46):	333.387
		Prob > LR:	0.000
McFadden's R2:	0.171	McFadden's Adj R2:	0.113
Maximum Likelihood R2:	0.651	Cragg & Uhler's R2:	0.652
AIC:	5.444	AIC*n:	1725.772
BIC:	114.458	BIC':	-68.478

Yes	.9953553	.4323304	2.30	0.021	.1480033	1.842707
friend						
Efficacy	.1666849	.5940316	0.28	0.779	-.9975957	1.330965
Other	1.357967	.9174727	1.48	0.139	-.440246	3.156181
natenroll_trans	-.5979159	.1759332	-3.40	0.001	-.9427386	-.2530931
natsitescat						
10-49 sites	-.1115708	.4287612	-0.26	0.795	-.9519273	.7287858
50-199 sites	-.042522	.5545389	-0.08	0.939	-1.129398	1.044354
200+ sites	-.211191	.7812864	-0.27	0.787	-1.742484	1.320102
Unknown	0	(empty)				
totalmo_trans	-.0040879	.2782841	-0.01	0.988	-.5495147	.5413389
moaccrdone_trans	.056276	.1084235	0.52	0.604	-.1562301	.2687821
precision						
Yes	.3504599	.5511804	0.64	0.525	-.7298338	1.430754
Conditional	0	(empty)				
disease_team						
0	0	(empty)				
1	1.691303	.9956054	1.70	0.089	-.2600474	3.642654
2	.1263508	.7228991	0.17	0.861	-1.290505	1.543207
4	3.162026	.9508872	3.33	0.001	1.298321	5.02573
5	-.6742233	.961524	-0.70	0.483	-2.558776	1.210329
6	0	(empty)				
7	-1.566581	1.210859	-1.29	0.196	-3.939821	.806658
8	0	(empty)				
9	0	(empty)				
10	1.401658	.6944057	2.02	0.044	.0406482	2.762669
11	-.3057761	1.346265	-0.23	0.820	-2.944407	2.332855
12	.35115	.7601954	0.46	0.644	-1.138806	1.841106
13	1.688676	.9896002	1.71	0.088	-.2509048	3.628257
14	1.04714	.7452352	1.41	0.160	-.4134945	2.507774
15	.0757538	1.072493	0.07	0.944	-2.026294	2.177802
16	1.951114	1.101122	1.77	0.076	-.2070452	4.109273
18	-.3888933	1.33963	-0.29	0.772	-3.01452	2.236733
19	0	(omitted)				
_cons	3.470272	2.553538	1.36	0.174	-1.53457	8.475115

Measures of Fit for logit of nonaccr

Log-Lik Intercept Only:	-240.960	Log-Lik Full Model:	-184.618
D(376):	369.236	LR(44):	112.684
		Prob > LR:	0.000
McFadden's R2:	0.234	McFadden's Adj R2:	-0.019
Maximum Likelihood R2:	0.227	Cragg & Uhler's R2:	0.340
McKelvey and Zavoina's R2:	0.421	Efron's R2:	0.252
Variance of y*:	5.685	Variance of error:	3.290
Count R2:	0.810	Adj Count R2:	0.210
AIC:	1.124	AIC*n:	491.236
BIC:	-1916.819	BIC':	154.833

Appendix F-9b: Site 560623 – Protocols per Staff Member and Protocols per FTE

Logistic regression	Number of obs	=	141
	LR chi2(34)	=	91.60
	Prob > chi2	=	0.0000
Log likelihood = -51.644642	Pseudo R2	=	0.4700

nonaccr	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
-----+-----					
prot_per_staff_cat					
2 < 2.5	2.803229	2.494603	1.12	0.261	-2.086104 7.692561
3.5 < 4	3.252908	2.236636	1.45	0.146	-1.130818 7.636633
4 < 4.5	-1.609136	2.313734	-0.70	0.487	-6.14397 2.925699
4.5 < 5	-.7275762	2.338514	-0.31	0.756	-5.310979 3.855827
5 < 5.5	1.891414	1.982967	0.95	0.340	-1.99513 5.777957
6 < 6.5	2.787865	2.134807	1.31	0.192	-1.39628 6.972009
9.5 < 10	1.129486	1.978953	0.57	0.568	-2.74919 5.008162
phase					
Phase 1/2	-.8668707	1.816399	-0.48	0.633	-4.426947 2.693206
Phase 2	-1.016128	1.942586	-0.52	0.601	-4.823526 2.79127
Phase 2/3	-1.854763	2.389797	-0.78	0.438	-6.538678 2.829153
Phase 3	-1.707225	2.065165	-0.83	0.408	-5.754873 2.340424
Phase 4	0	(empty)			
None	-5.464125	3.070243	-1.78	0.075	-11.48169 .5534396
irb					
Local	.9444075	.8593266	1.10	0.272	-.7398416 2.628657
sponsor_type					
Industry	2.905521	2.117167	1.37	0.170	-1.24405 7.055091

National Group	4.253149	2.224887	1.91	0.056	-.1075483	8.613847
random						
Yes	.4273019	1.001688	0.43	0.670	-1.535971	2.390575
friend						
Efficacy	.0121325	2.03356	0.01	0.995	-3.973572	3.997837
Other	0	(empty)				
natenroll_trans	-.7349452	.4980628	-1.48	0.140	-1.71113	.24124
natsitescat						
10-49 sites	.8073229	1.520934	0.53	0.596	-2.173653	3.788298
50-199 sites	2.048945	1.540358	1.33	0.183	-.9701007	5.067991
200+ sites	1.74536	1.765929	0.99	0.323	-1.715797	5.206517
Unknown	0	(empty)				
totalmo_trans	-1.887766	.860708	-2.19	0.028	-3.574723	-.2008094
moaccrdone_trans	1.425972	.3668929	3.89	0.000	.7068753	2.145069
primary_purpose						
Supportive Care	1.818952	1.338051	1.36	0.174	-.8035791	4.441483
precision						
Yes	1.651853	2.353038	0.70	0.483	-2.960017	6.263723
Conditional	1.644091	1.970372	0.83	0.404	-2.217767	5.505949
disease_team						
1	-1.610167	1.193159	-1.35	0.177	-3.948716	.7283814
2	-1.89162	2.021701	-0.94	0.349	-5.854081	2.07084
3	-.1219478	1.299511	-0.09	0.925	-2.668942	2.425046
4	.2352033	1.205688	0.20	0.845	-2.127902	2.598308
5	-2.500521	1.543771	-1.62	0.105	-5.526257	.525215
6	-.229382	1.332362	-0.17	0.863	-2.840763	2.381999
7	-1.230279	1.243187	-0.99	0.322	-3.666881	1.206323
9	-2.9432	2.391098	-1.23	0.218	-7.629666	1.743265
10	0	(empty)				
_cons	.5102064	4.377455	0.12	0.907	-8.069448	9.089861

Measures of Fit for logit of nonaccr

Log-Lik Intercept Only:	-97.446	Log-Lik Full Model:	-51.645
D(92):	103.289	LR(34):	91.603
		Prob > LR:	0.000
McFadden's R2:	0.470	McFadden's Adj R2:	-0.033
Maximum Likelihood R2:	0.478	Cragg & Uhler's R2:	0.638
McKelvey and Zavoina's R2:	0.784	Efron's R2:	0.548
Variance of y*:	15.214	Variance of error:	3.290
Count R2:	0.851	Adj Count R2:	0.682
AIC:	1.428	AIC*n:	201.289
BIC:	-351.997	BIC':	76.654

Appendix F-9c: Site 602591 – Protocols per Staff Member

Logistic regression	Number of obs	=	320
	LR chi2(39)	=	102.58
	Prob > chi2	=	0.0000
Log likelihood = -138.83176	Pseudo R2	=	0.2698

nonaccr	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]

prot_per_staff_cat					
< 1	0	(empty)			
1 < 1.5	4.177328	1.877199	2.23	0.026	.4980855 7.85657
1.5 < 2	2.569165	1.027154	2.50	0.012	.5559809 4.582349
2 < 2.5	3.10822	1.487802	2.09	0.037	.1921826 6.024258
2.5 < 3	1.102384	1.113392	0.99	0.322	-1.079823 3.284592
3 < 3.5	1.490755	.5711595	2.61	0.009	.3713026 2.610207
3.5 < 4	.2512912	.8231716	0.31	0.760	-1.362095 1.864678
4 < 4.5	0	(empty)			
4.5 < 5	.3621527	.5712216	0.63	0.526	-.757421 1.481726
5 < 5.5	.7887872	.5680721	1.39	0.165	-.3246136 1.902188
5.5 < 6	.910271	.8699079	1.05	0.295	-.7947171 2.615259
6 < 6.5	-.3423836	.7898139	-0.43	0.665	-1.89039 1.205623
6.5 < 7	0	(omitted)			

phase					
Phase 1/2	1.26365	.7160531	1.76	0.078	-.1397879 2.667089
Phase 2	.6704224	.6294466	1.07	0.287	-.5632702 1.904115
Phase 2/3	2.812668	1.610117	1.75	0.081	-.3431026 5.968438
Phase 3	1.420546	.8088726	1.76	0.079	-.1648154 3.005907
Phase 4	0	(empty)			
None	0	(empty)			

sponsor_type						
Instititutional	1.953236	1.546731	1.26	0.207	-1.078301	4.984774
Industry	2.798414	1.406075	1.99	0.047	.0425585	5.55427
National Group	2.766785	1.434257	1.93	0.054	-.0443058	5.577877
peds						
Yes	-2.029442	1.817538	-1.12	0.264	-5.591752	1.532868
natenroll_trans	-2.528338	.6300088	-4.01	0.000	-3.763133	-1.293543
natsitescat						
10-49 sites	.5656422	.5475621	1.03	0.302	-.5075598	1.638844
50-199 sites	.5421662	.6839758	0.79	0.428	-.7984018	1.882734
200+ sites	1.116639	.8356738	1.34	0.181	-.5212514	2.75453
Unknown	-.9076104	1.464821	-0.62	0.536	-3.778607	1.963386
totalmo_trans	-1.564294	.9098123	-1.72	0.086	-3.347494	.2189052
moaccrdone_trans	.4331212	.1353168	3.20	0.001	.1679053	.6983372
disease_team						
1	-.0327439	1.286516	-0.03	0.980	-2.55427	2.488782
2	-1.468873	1.961368	-0.75	0.454	-5.313084	2.375338
3	-1.242617	1.257325	-0.99	0.323	-3.706928	1.221694
4	.4135485	1.288463	0.32	0.748	-2.111792	2.938889
5	.3159367	1.275261	0.25	0.804	-2.183529	2.815403
6	-1.403994	1.678282	-0.84	0.403	-4.693366	1.885378
7	-.8141893	1.212617	-0.67	0.502	-3.190876	1.562497
8	-.1696348	1.291001	-0.13	0.895	-2.699951	2.360682
9	-.0268278	1.213914	-0.02	0.982	-2.406055	2.352399
10	-.9635596	1.433206	-0.67	0.501	-3.772592	1.845473
11	-1.869104	1.395206	-1.34	0.180	-4.603658	.8654507
12	-.662139	1.357523	-0.49	0.626	-3.322836	1.998558
13	1.388227	1.349518	1.03	0.304	-1.256781	4.033234
natenroll_trans	0	(omitted)				
totalmo_trans	0	(omitted)				
c.natenroll_trans#c.totalmo_trans	.4835229	.1788428	2.70	0.007	.1329974	.8340484
_cons	2.527652	3.52291	0.72	0.473	-4.377125	9.432428

Measures of Fit for logit of nonaccr

Log-Lik Intercept Only:	-190.122	Log-Lik Full Model:	-138.832
D(268):	277.664	LR(39):	102.580
		Prob > LR:	0.000
McFadden's R2:	0.270	McFadden's Adj R2:	-0.004
Maximum Likelihood R2:	0.274	Cragg & Uhler's R2:	0.394
McKelvey and Zavoina's R2:	0.484	Efron's R2:	0.301
Variance of y*:	6.372	Variance of error:	3.290
Count R2:	0.778	Adj Count R2:	0.211
AIC:	1.193	AIC*n:	381.664
BIC:	-1268.247	BIC':	122.385

Appendix F-9d: Site 602591 – Protocols per FTE

Logistic regression	Number of obs	=	321
	LR chi2(39)	=	104.90
	Prob > chi2	=	0.0000
Log likelihood = -138.93842	Pseudo R2	=	0.2740

nonaccr	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]

prot_per_fte_cat					
< 1	0	(empty)			
1 < 1.5	4.377099	1.923237	2.28	0.023	.6076246 8.146573
1.5 < 2	2.785909	1.103307	2.53	0.012	.623466 4.948352
2 < 2.5	3.288392	1.530922	2.15	0.032	.2878402 6.288944
2.5 < 3	1.333391	1.187577	1.12	0.262	-.9942172 3.661
3 < 3.5	1.678577	.702042	2.39	0.017	.3025997 3.054554
3.5 < 4	2.861418	1.657439	1.73	0.084	-.3871032 6.109938
4 < 4.5	.607281	.8792321	0.69	0.490	-1.115982 2.330544
5 < 5.5	.565333	.6765283	0.84	0.403	-.7606382 1.891304
5.5 < 6	1.021173	.6628443	1.54	0.123	-.2779783 2.320324
7 < 7.5	.193315	.742587	0.26	0.795	-1.262129 1.648759
7.5 < 8	0	(omitted)			
phase					
Phase 1/2	1.312959	.7025668	1.87	0.062	-.0640463 2.689965
Phase 2	.6736871	.6195758	1.09	0.277	-.540659 1.888033
Phase 2/3	2.988829	1.624049	1.84	0.066	-.1942489 6.171907
Phase 3	1.468747	.8058668	1.82	0.068	-.1107225 3.048217
Phase 4	0	(empty)			
None	0	(empty)			

sponsor_type						
Instititutional	1.855027	1.55212	1.20	0.232	-1.187071	4.897126
Industry	2.86453	1.41697	2.02	0.043	.0873207	5.641739
National Group	2.804641	1.447906	1.94	0.053	-.0332027	5.642485
peds						
Yes	-1.95366	1.741942	-1.12	0.262	-5.367803	1.460482
natenroll_trans	-2.524316	.6259179	-4.03	0.000	-3.751093	-1.29754
natsitescat						
10-49 sites	.5198532	.539364	0.96	0.335	-.5372808	1.576987
50-199 sites	.5318709	.6782135	0.78	0.433	-.7974032	1.861145
200+ sites	1.07484	.8318462	1.29	0.196	-.5555484	2.705229
Unknown	-.8608639	1.421532	-0.61	0.545	-3.647015	1.925288
totalmo_trans						
totalmo_trans	-1.518417	.8962975	-1.69	0.090	-3.275128	.2382936
moaccrdone_trans	.4338303	.1345316	3.22	0.001	.1701531	.6975075
disease_team						
1	.0028069	1.271236	0.00	0.998	-2.488769	2.494383
2	-1.413248	1.956631	-0.72	0.470	-5.248173	2.421678
3	-1.316401	1.247588	-1.06	0.291	-3.761628	1.128826
4	.5185478	1.253871	0.41	0.679	-1.938995	2.976091
5	.3902721	1.264358	0.31	0.758	-2.087825	2.868369
6	-1.90912	1.829294	-1.04	0.297	-5.49447	1.676231
7	-.8439275	1.185877	-0.71	0.477	-3.168204	1.480349
8	-.1513297	1.269591	-0.12	0.905	-2.639682	2.337023
9	.0001617	1.20582	0.00	1.000	-2.363202	2.363526
10	-.8774738	1.428696	-0.61	0.539	-3.677667	1.92272
11	-1.848915	1.376224	-1.34	0.179	-4.546264	.8484333
12	-.6308718	1.352445	-0.47	0.641	-3.281615	2.019872
13	1.466281	1.321344	1.11	0.267	-1.123506	4.056067
c.natenroll_trans#c.totalmo_trans						
c.natenroll_trans#c.totalmo_trans	.4793661	.1774882	2.70	0.007	.1314955	.8272366
_cons						
_cons	2.132353	3.473111	0.61	0.539	-4.674819	8.939526

Measures of Fit for logit of nonaccr

Log-Lik Intercept Only:	-191.386	Log-Lik Full Model:	-138.938
D(270):	277.877	LR(39):	104.895
		Prob > LR:	0.000
McFadden's R2:	0.274	McFadden's Adj R2:	0.008
Maximum Likelihood R2:	0.279	Cragg & Uhler's R2:	0.400
McKelvey and Zavoina's R2:	0.490	Efron's R2:	0.303
Variance of y*:	6.450	Variance of error:	3.290
Count R2:	0.769	Adj Count R2:	0.187
AIC:	1.183	AIC*n:	379.877
BIC:	-1280.412	BIC':	120.191

Yes	-.3064246	.49096	-0.62	0.533	-1.268688	.6558393
random						
Yes	-.548951	.3598507	-1.53	0.127	-1.254245	.1563435
placebo						
Yes	-.5638534	.4457912	-1.26	0.206	-1.437588	.3098813
friend						
Efficacy	.5993692	.5539131	1.08	0.279	-.4862805	1.685019
Other	-.965603	.9252199	-1.04	0.297	-2.779001	.8477946
natenroll_trans	-.820999	.9010449	-0.91	0.362	-2.587015	.9450165
natsitescat						
10-49 sites	-.4137136	.4020796	-1.03	0.304	-1.201775	.3743479
50-199 sites	-.1177223	.564229	-0.21	0.835	-1.223591	.9881462
200+ sites	-.5276158	.7712703	-0.68	0.494	-2.039278	.9840462
Unknown	.1552869	1.175626	0.13	0.895	-2.148898	2.459472
totalmo_trans	.0340255	.2662108	0.13	0.898	-.4877379	.555789
moaccrdone_trans	-.3090754	.1069511	-2.89	0.004	-.5186957	-.0994551
primary_purpose						
Supportive Care	1.7547	.7839825	2.24	0.025	.2181222	3.291277
precision						
Yes	-.6030446	.5690074	-1.06	0.289	-1.718279	.5121894
Conditional	-.2499515	1.175015	-0.21	0.832	-2.552939	2.053036
disease_team						
0	0	(empty)				
1	-1.380451	.8728092	-1.58	0.114	-3.091126	.3302237
2	.0977446	.5927556	0.16	0.869	-1.064035	1.259524
4	-1.825697	.9552696	-1.91	0.056	-3.697991	.0465969
5	.8328755	.800763	1.04	0.298	-.7365911	2.402342
6	0	(empty)				
7	-.0092184	.6683914	-0.01	0.989	-1.319242	1.300805
8	1.463788	1.038141	1.41	0.159	-.570931	3.498508
9	.9394294	.8559893	1.10	0.272	-.7382788	2.617138
10	-.3643794	.6087724	-0.60	0.549	-1.557551	.8287925
11	-.0076747	.9208846	-0.01	0.993	-1.812575	1.797226
12	.0026787	.6526606	0.00	0.997	-1.276513	1.28187
13	-1.710646	1.288556	-1.33	0.184	-4.23617	.814877
14	-.9893564	.6613969	-1.50	0.135	-2.285671	.3069578

15		-.4406751	.9014779	-0.49	0.625	-2.207539	1.326189
16		-1.22004	1.319403	-0.92	0.355	-3.806022	1.365943
18		.2134486	1.122469	0.19	0.849	-1.98655	2.413447
19		0	(omitted)				
natenroll_trans		0	(omitted)				
sponsor_type#c.natenroll_trans							
Institutional		1.572882	.9646593	1.63	0.103	-.3178154	3.46358
Industry		1.527649	.8998681	1.70	0.090	-.2360602	3.291358
National Group		2.544008	.9240907	2.75	0.006	.7328233	4.355192
_cons		5.602472	4.991218	1.12	0.262	-4.180136	15.38508

Measures of Fit for logit of actualaccrual_binary

Log-Lik Intercept Only:	-319.473	Log-Lik Full Model:	-216.241
D(401):	432.482	LR(53):	206.463
		Prob > LR:	0.000
McFadden's R2:	0.323	McFadden's Adj R2:	0.104
Maximum Likelihood R2:	0.355	Cragg & Uhler's R2:	0.478
McKelvey and Zavoina's R2:	0.545	Efron's R2:	0.388
Variance of y*:	7.223	Variance of error:	3.290
Count R2:	0.786	Adj Count R2:	0.482
AIC:	1.215	AIC*n:	572.482
BIC:	-2035.616	BIC':	119.744

Appendix F-10b: Site 512786 – Protocols per Staff Member and Protocols per FTE

Logistic regression	Number of obs	=	391
	LR chi2(36)	=	101.78
	Prob > chi2	=	0.0000
Log likelihood = -205.30026	Pseudo R2	=	0.1986

actualaccrual_binary	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]		
-----+-----							
prot_per_staff_cat							
1.5 < 2		-.9336736	1.42246	-0.66	0.512	-3.721645	1.854298
2 < 2.5		1.724327	1.29687	1.33	0.184	-.817492	4.266145
2.5 < 3		1.179687	1.291234	0.91	0.361	-1.351085	3.710458
3 < 3.5		1.243866	1.270387	0.98	0.328	-1.246046	3.733778
3.5 < 4		1.960542	1.36315	1.44	0.150	-.7111827	4.632267
5 < 5.5		1.043684	1.278461	0.82	0.414	-1.462053	3.54942

phase						
Phase 1	.7857351	.8891861	0.88	0.377	-.9570376	2.528508
Phase 1/2	1.517308	.8992131	1.69	0.092	-.245117	3.279733
Phase 2	.3825153	.7971708	0.48	0.631	-1.179911	1.944941
Phase 2/3	0	(empty)				
Phase 3	.2464939	.9453868	0.26	0.794	-1.60643	2.099418
Phase 4	0	(empty)				
irb						
Local	-.2474555	.3207753	-0.77	0.440	-.8761636	.3812526
sponsor_type						
Instituitonal	.5738782	.6575942	0.87	0.383	-.7149827	1.862739
Industry	-.7023929	.5190743	-1.35	0.176	-1.71976	.3149741
National Group	.0147681	.6654317	0.02	0.982	-1.289454	1.31899
random						
Yes	-.2999222	.3916158	-0.77	0.444	-1.067475	.4676306
placebo						
Yes	.7373709	.4620877	1.60	0.111	-.1683044	1.643046
priend						
Efficacy	0	(omitted)				
natenroll_trans	.3627562	.1615876	2.24	0.025	.0460504	.679462
natsitescat						
10-49 sites	-.0410987	.36825	-0.11	0.911	-.7628555	.6806581
50-199 sites	.1833098	.5183173	0.35	0.724	-.8325735	1.199193
200+ sites	-1.081239	.7829627	-1.38	0.167	-2.615817	.4533399
Unknown	.5690631	.5827743	0.98	0.329	-.5731534	1.71128
moaccrdone_trans	-.4489569	.08942	-5.02	0.000	-.6242168	-.273697
disease_team						
2	.4462578	.8167767	0.55	0.585	-1.154595	2.047111
3	1.024697	.6560782	1.56	0.118	-.2611926	2.310587
4	-.1749001	.6362556	-0.27	0.783	-1.421938	1.072138
5	-.3083736	.6980528	-0.44	0.659	-1.676532	1.059785
6	.5261692	.8645721	0.61	0.543	-1.168361	2.220699
7	1.720281	1.267952	1.36	0.175	-.7648589	4.205421
8	0	(empty)				
9	.5937316	1.581515	0.38	0.707	-2.50598	3.693443
10	.1512041	.5947723	0.25	0.799	-1.014528	1.316936

12		-.2275248	.6958465	-0.33	0.744	-1.591359	1.136309
13		.3322806	.742734	0.45	0.655	-1.123451	1.788013
14		-1.404915	.7613492	-1.85	0.065	-2.897132	.0873023
15		.5022325	.8074895	0.62	0.534	-1.080418	2.084883
16		-.3643766	1.521797	-0.24	0.811	-3.347045	2.618291
17		.8484082	.6652848	1.28	0.202	-.4555261	2.152343
_cons		-1.356284	1.748071	-0.78	0.438	-4.78244	2.069872

Measures of Fit for logit of actualaccrual_binary

Log-Lik Intercept Only:	-256.191	Log-Lik Full Model:	-205.300
D(341):	410.601	LR(36):	101.782
		Prob > LR:	0.000
McFadden's R2:	0.199	McFadden's Adj R2:	0.003
Maximum Likelihood R2:	0.229	Cragg & Uhler's R2:	0.314
McKelvey and Zavoina's R2:	0.331	Efron's R2:	0.243
Variance of y*:	4.914	Variance of error:	3.290
Count R2:	0.749	Adj Count R2:	0.310
AIC:	1.306	AIC*n:	510.601
BIC:	-1624.729	BIC':	113.091

Appendix F-11 Site-Specific Full Models Assessing the Impact of Protocol Workload on Activation Time

Appendix F-11a: Site 494048 – Protocols per Staff Member and Protocols per FTE

Source	SS	df	MS	Number of obs	=	477
				F(52, 424)	=	6.78
Model	3410.54604	52	65.5874238	Prob > F	=	0.0000
Residual	4102.45694	424	9.675606	R-squared	=	0.4540
				Adj R-squared	=	0.3870
Total	7513.00298	476	15.7836197	Root MSE	=	3.1106

activationtime_trans	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
prot_per_fte_cat						
1 < 1.5	.6401362	2.580909	0.25	0.804	-4.432833 5.713105	
2.5 < 3	1.267103	2.469081	0.51	0.608	-3.586059 6.120265	
3 < 3.5	3.874678	2.432947	1.59	0.112	-.9074617 8.656817	
4 < 4.5	4.320714	2.355553	1.83	0.067	-.3093019 8.950729	
4.5 < 5	3.531556	2.34799	1.50	0.133	-1.083594 8.146705	
5 < 5.5	4.965227	2.313018	2.15	0.032	.4188163 9.511637	
5.5 < 6	4.918804	2.379233	2.07	0.039	.2422438 9.595364	
6 < 6.5	5.617226	2.396211	2.34	0.020	.9072934 10.32716	
7 < 7.5	4.938637	2.359804	2.09	0.037	.3002658 9.577009	
7.5 < 8	5.591832	2.362517	2.37	0.018	.9481293 10.23554	
8 < 8.5	4.469845	2.367985	1.89	0.060	-.1846061 9.124297	
phase						
Phase 1	4.032155	3.241842	1.24	0.214	-2.339928 10.40424	
Phase 1/2	4.43808	3.237559	1.37	0.171	-1.925583 10.80174	
Phase 2	3.89181	3.208716	1.21	0.226	-2.41516 10.19878	
Phase 2/3	3.882719	3.531435	1.10	0.272	-3.05858 10.82402	
Phase 3	4.816294	3.240742	1.49	0.138	-1.553626 11.18621	
Phase 4	-.396709	3.465643	-0.11	0.909	-7.208689 6.415271	
None	4.827927	3.349648	1.44	0.150	-1.756056 11.41191	
sponsor_type						
Instititutional	-1.074325	1.042853	-1.03	0.304	-3.124131 .9754805	
Industry	-1.785521	1.001811	-1.78	0.075	-3.754656 .1836133	
National Group	-4.949399	1.072825	-4.61	0.000	-7.058116 -2.840682	
peds						
Yes	-.8277635	.5598274	-1.48	0.140	-1.928146 .2726191	

placebo							
Yes	.2456795	.502844	0.49	0.625	-.742698	1.234057	
friend							
Efficacy	-.0169606	.6390636	-0.03	0.979	-1.273088	1.239167	
Other	-1.567532	1.156049	-1.36	0.176	-3.839833	.7047685	
natenroll_trans	-.6182098	.2505267	-2.47	0.014	-1.110639	-.1257809	
natsitescat							
10-49 sites	-.4851024	.4579091	-1.06	0.290	-1.385157	.4149522	
50-199 sites	-.886734	.6182483	-1.43	0.152	-2.101947	.3284793	
200+ sites	-.0946397	.8626157	-0.11	0.913	-1.790175	1.600896	
Unknown	-2.1508	1.327923	-1.62	0.106	-4.760932	.4593315	
totalmo_trans	-.1249586	.3073752	-0.41	0.685	-.7291274	.4792103	
moaccrdone_trans	-.500623	.4135416	-1.21	0.227	-1.31347	.3122239	
precision							
Yes	.1788624	.6390941	0.28	0.780	-1.077325	1.43505	
Conditional	2.946083	1.459517	2.02	0.044	.0772936	5.814873	
disease_team							
1	3.698118	1.925475	1.92	0.055	-.0865481	7.482783	
2	3.304	1.85046	1.79	0.075	-.3332168	6.941217	
4	6.433535	2.022345	3.18	0.002	2.458466	10.4086	
5	2.774984	1.930291	1.44	0.151	-1.019147	6.569114	
6	3.275894	3.88432	0.84	0.400	-4.359027	10.91082	
7	4.442422	1.903279	2.33	0.020	.7013843	8.18346	
8	6.201632	2.18841	2.83	0.005	1.900148	10.50312	
9	4.484011	1.961981	2.29	0.023	.6275919	8.340431	
10	4.430418	1.888445	2.35	0.019	.7185375	8.142298	
11	7.034539	2.106587	3.34	0.001	2.893886	11.17519	
12	4.398516	1.895484	2.32	0.021	.6728008	8.12423	
13	1.591567	2.091017	0.76	0.447	-2.518483	5.701616	
14	3.548569	1.902008	1.87	0.063	-.1899692	7.287107	
15	4.258254	2.037018	2.09	0.037	.2543435	8.262165	
16	5.729309	2.260303	2.53	0.012	1.286515	10.1721	
18	3.738199	2.152048	1.74	0.083	-.4918116	7.968209	
19	4.26431	1.906366	2.24	0.026	.5172056	8.011415	
natenroll_trans	0	(omitted)					
moaccrdone_trans	0	(omitted)					
c.natenroll_trans#c.moaccrdone_trans	.2934644	.0839637	3.50	0.001	.1284276	.4585013	

_cons	2.052021	4.596599	0.45	0.656	-6.982938	11.08698
-------	----------	----------	------	-------	-----------	----------

Measures of Fit for regress of activationtime_trans

Log-Lik Intercept Only:	-1334.348	Log-Lik Full Model:	-1190.044
D(413):	2380.088	LR(52):	288.608
		Prob > LR:	0.000
R2:	0.454	Adjusted R2:	0.387
AIC:	5.258	AIC*n:	2508.088
BIC:	-167.097	BIC':	32.102

Appendix F-11b: Site 602591 – Protocols per Staff Member

Source	SS	df	MS	Number of obs	=	329
Model	4121.04395	44	93.6600898	F(44, 284)	=	11.67
Residual	2278.35867	284	8.02238969	Prob > F	=	0.0000
Total	6399.40262	328	19.5103738	R-squared	=	0.6440
				Adj R-squared	=	0.5888
				Root MSE	=	2.8324

activationtime_t~s	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
prot_per_staff_cat						
1 < 1.5	1.82532	3.227378	0.57	0.572	-4.527296	8.177936
1.5 < 2	1.071134	3.81389	0.28	0.779	-6.435945	8.578214
2 < 2.5	5.595936	3.941469	1.42	0.157	-2.162263	13.35414
2.5 < 3	6.044761	3.803427	1.59	0.113	-1.441724	13.53125
3 < 3.5	3.684396	3.745935	0.98	0.326	-3.688923	11.05772
3.5 < 4	4.306696	3.758691	1.15	0.253	-3.091732	11.70512
4 < 4.5	14.95096	4.717664	3.17	0.002	5.664936	24.23698
4.5 < 5	4.133627	3.715836	1.11	0.267	-3.180446	11.4477
5 < 5.5	5.448241	3.714735	1.47	0.144	-1.863665	12.76015
5.5 < 6	8.446239	3.818954	2.21	0.028	.9291933	15.96328
6 < 6.5	4.386137	3.738046	1.17	0.242	-2.971653	11.74393
6.5 < 7	4.754915	3.73683	1.27	0.204	-2.600482	12.11031
phase						
Phase 1/2	1.577493	.7486267	2.11	0.036	.1039324	3.051054
Phase 2	-.0677652	.6205892	-0.11	0.913	-1.289303	1.153773
Phase 2/3	3.047267	2.147579	1.42	0.157	-1.179925	7.274458
Phase 3	1.349714	.8276263	1.63	0.104	-.2793457	2.978774
Phase 4	-3.667035	2.978205	-1.23	0.219	-9.529192	2.195121

None	-.9686456	1.756837	-0.55	0.582	-4.42672	2.489428
irb						
Local	4.176806	3.680562	1.13	0.257	-3.067836	11.42145
sponsor_type						
Instituitonal	-2.313041	1.099585	-2.10	0.036	-4.477412	-.1486702
Industry	-3.458079	.9939931	-3.48	0.001	-5.414608	-1.501551
National Group	-7.835571	1.051393	-7.45	0.000	-9.905083	-5.766058
natenroll_trans	-.4239394	.1871229	-2.27	0.024	-.7922631	-.0556157
natsitescat						
10-49 sites	.2733491	.5536683	0.49	0.622	-.8164651	1.363163
50-199 sites	1.007054	.6978125	1.44	0.150	-.3664863	2.380595
200+ sites	.3196123	.8845464	0.36	0.718	-1.421487	2.060711
Unknown	1.59025	1.050892	1.51	0.131	-.4782753	3.658775
moaccrdone_trans	.6535736	.1317037	4.96	0.000	.3943343	.9128128
totalmo_trans	.4176803	.3692318	1.13	0.259	-.309098	1.144459
precision						
Yes	-.135706	.9252577	-0.15	0.883	-1.956939	1.685527
Conditional	1.353011	1.169849	1.16	0.248	-.9496644	3.655687
disease_team						
1	2.238143	1.632924	1.37	0.172	-.9760263	5.452312
2	-.0160432	2.425657	-0.01	0.995	-4.79059	4.758503
3	-1.184485	1.562832	-0.76	0.449	-4.260688	1.891718
4	.1908312	1.610235	0.12	0.906	-2.978678	3.360341
5	-3.043364	1.626562	-1.87	0.062	-6.245011	.1582838
6	-1.446972	1.838255	-0.79	0.432	-5.065305	2.171361
7	.55011	1.556548	0.35	0.724	-2.513725	3.613945
8	.7228869	1.633614	0.44	0.658	-2.49264	3.938414
9	.0976557	1.55785	0.06	0.950	-2.968742	3.164053
10	-.705667	1.700543	-0.41	0.678	-4.052935	2.641601
11	.1393758	1.614539	0.09	0.931	-3.038605	3.317356
12	-.2866143	1.634594	-0.18	0.861	-3.504072	2.930843
13	-.2744346	1.666048	-0.16	0.869	-3.553803	3.004934
_cons	8.037052	3.246114	2.48	0.014	1.647557	14.42655

Measures of Fit for regress of activationtime_trans

Log-Lik Intercept Only:	-955.051	Log-Lik Full Model:	-785.163
D(277):	1570.327	LR(44):	339.775
		Prob > LR:	0.000
R2:	0.644	Adjusted R2:	0.589
AIC:	5.089	AIC*n:	1674.327
BIC:	-35.181	BIC':	-84.748

Appendix F-11c: Site 602591 – Protocols per FTE

-----+-----				F(44, 284)	=	11.67
Model	4121.04395	44	93.6600898	Prob > F	=	0.0000
Residual	2278.35867	284	8.02238969	R-squared	=	0.6440
-----+-----				Adj R-squared	=	0.5888
Total	6399.40262	328	19.5103738	Root MSE	=	2.8324

activationtime_t~s	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
-----+-----						
prot_per_staff_cat						
1 < 1.5	1.82532	3.227378	0.57	0.572	-4.527296	8.177936
1.5 < 2	1.071134	3.81389	0.28	0.779	-6.435945	8.578214
2 < 2.5	5.595936	3.941469	1.42	0.157	-2.162263	13.35414
2.5 < 3	6.044761	3.803427	1.59	0.113	-1.441724	13.53125
3 < 3.5	3.684396	3.745935	0.98	0.326	-3.688923	11.05772
3.5 < 4	4.306696	3.758691	1.15	0.253	-3.091732	11.70512
4 < 4.5	14.95096	4.717664	3.17	0.002	5.664936	24.23698
4.5 < 5	4.133627	3.715836	1.11	0.267	-3.180446	11.4477
5 < 5.5	5.448241	3.714735	1.47	0.144	-1.863665	12.76015
5.5 < 6	8.446239	3.818954	2.21	0.028	.9291933	15.96328
6 < 6.5	4.386137	3.738046	1.17	0.242	-2.971653	11.74393
6.5 < 7	4.754915	3.73683	1.27	0.204	-2.600482	12.11031
phase						
Phase 1/2	1.577493	.7486267	2.11	0.036	.1039324	3.051054
Phase 2	-.0677652	.6205892	-0.11	0.913	-1.289303	1.153773
Phase 2/3	3.047267	2.147579	1.42	0.157	-1.179925	7.274458
Phase 3	1.349714	.8276263	1.63	0.104	-.2793457	2.978774
Phase 4	-3.667035	2.978205	-1.23	0.219	-9.529192	2.195121
None	-.9686456	1.756837	-0.55	0.582	-4.42672	2.489428
irb						
Local	4.176806	3.680562	1.13	0.257	-3.067836	11.42145

sponsor_type						
Instititutional	-2.313041	1.099585	-2.10	0.036	-4.477412	-.1486702
Industry	-3.458079	.9939931	-3.48	0.001	-5.414608	-1.501551
National Group	-7.835571	1.051393	-7.45	0.000	-9.905083	-5.766058
natsitescat						
10-49 sites	.2733491	.5536683	0.49	0.622	-.8164651	1.363163
50-199 sites	1.007054	.6978125	1.44	0.150	-.3664863	2.380595
200+ sites	.3196123	.8845464	0.36	0.718	-1.421487	2.060711
Unknown	1.59025	1.050892	1.51	0.131	-.4782753	3.658775
natenroll_trans						
totalmo_trans	-.4239394	.1871229	-2.27	0.024	-.7922631	-.0556157
moaccrdone_trans	.4176803	.3692318	1.13	0.259	-.309098	1.144459
precision						
Yes	.6535736	.1317037	4.96	0.000	.3943343	.9128128
Conditional	-.135706	.9252577	-0.15	0.883	-1.956939	1.685527
	1.353011	1.169849	1.16	0.248	-.9496644	3.655687
disease_team						
1	2.238143	1.632924	1.37	0.172	-.9760263	5.452312
2	-.0160432	2.425657	-0.01	0.995	-4.79059	4.758503
3	-1.184485	1.562832	-0.76	0.449	-4.260688	1.891718
4	.1908312	1.610235	0.12	0.906	-2.978678	3.360341
5	-3.043364	1.626562	-1.87	0.062	-6.245011	.1582838
6	-1.446972	1.838255	-0.79	0.432	-5.065305	2.171361
7	.55011	1.556548	0.35	0.724	-2.513725	3.613945
8	.7228869	1.633614	0.44	0.658	-2.49264	3.938414
9	.0976557	1.55785	0.06	0.950	-2.968742	3.164053
10	-.705667	1.700543	-0.41	0.678	-4.052935	2.641601
11	.1393758	1.614539	0.09	0.931	-3.038605	3.317356
12	-.2866143	1.634594	-0.18	0.861	-3.504072	2.930843
13	-.2744346	1.666048	-0.16	0.869	-3.553803	3.004934
_cons	8.037052	3.246114	2.48	0.014	1.647557	14.42655

Measures of Fit for regress of activationtime_trans

Log-Lik Intercept Only:	-955.051	Log-Lik Full Model:	-785.163
D(277):	1570.327	LR(44):	339.775
		Prob > LR:	0.000
R2:	0.644	Adjusted R2:	0.589
AIC:	5.089	AIC*n:	1674.327
BIC:	-35.181	BIC':	-84.748

Appendix F-11d: Site 696337 – Protocols per Staff Member and Protocols per FTE

Source	SS	df	MS	Number of obs	=	329
Model	3959.98834	43	92.092752	F(43, 285)	=	10.76
Residual	2439.41429	285	8.55934838	Prob > F	=	0.0000
				R-squared	=	0.6188
				Adj R-squared	=	0.5613
Total	6399.40262	328	19.5103738	Root MSE	=	2.9256

activationtime~s	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
prot_per_fte_cat						
1 < 1.5	1.579938	3.33216	0.47	0.636	-4.978828	8.138704
1.5 < 2	.9061065	3.936274	0.23	0.818	-6.84175	8.653963
2 < 2.5	5.144736	4.065759	1.27	0.207	-2.857989	13.14746
2.5 < 3	5.665606	3.924682	1.44	0.150	-2.059434	13.39065
3 < 3.5	3.30052	3.865879	0.85	0.394	-4.308776	10.90982
3.5 < 4	4.046896	4.170961	0.97	0.333	-4.162901	12.25669
4 < 4.5	4.694175	3.874745	1.21	0.227	-2.932574	12.32092
5 < 5.5	3.885481	3.835104	1.01	0.312	-3.663241	11.4342
5.5 < 6	5.798206	3.835262	1.51	0.132	-1.750828	13.34724
7 < 7.5	4.130418	3.84975	1.07	0.284	-3.447132	11.70797
7.5 < 8	5.065087	3.869712	1.31	0.192	-2.551755	12.68193
phase						
Phase 1/2	1.630534	.7623483	2.14	0.033	.129987	3.131082
Phase 2	-.3777407	.6334675	-0.60	0.551	-1.624609	.8691276
Phase 2/3	2.690354	2.223118	1.21	0.227	-1.685459	7.066167
Phase 3	.9949833	.8538308	1.17	0.245	-.6856312	2.675598
Phase 4	-4.052736	3.070817	-1.32	0.188	-10.09709	1.991622
None	-1.248899	1.819398	-0.69	0.493	-4.830061	2.332262
irb						
Local	4.796707	3.793809	1.26	0.207	-2.670733	12.26415
sponsor_type						
Institutitonal	-2.204228	1.135801	-1.94	0.053	-4.439851	.0313958
Industry	-3.215659	1.026796	-3.13	0.002	-5.236725	-1.194594
National Group	-7.799201	1.09114	-7.15	0.000	-9.946916	-5.651487
natenroll_trans	-.3796699	.1924427	-1.97	0.049	-.7584592	-.0008807
natsitescat						
10-49 sites	.2673926	.5651228	0.47	0.636	-.8449513	1.379736

50-199 sites		1.152432	.7171181	1.61	0.109	-.2590882	2.563951
200+ sites		.3433708	.913766	0.38	0.707	-1.455216	2.141957
Unknown		1.493045	1.084202	1.38	0.170	-.6410136	3.627103
totalmo_trans		.3237807	.3795643	0.85	0.394	-.4233242	1.070886
moaccrdone_trans		.7485384	.1352247	5.54	0.000	.4823724	1.014704
precision							
Yes		.2116883	.9477448	0.22	0.823	-1.653779	2.077156
Conditional		1.613906	1.205321	1.34	0.182	-.7585546	3.986366
disease_team							
1		1.556161	1.670655	0.93	0.352	-1.732227	4.844549
2		-.4072158	2.498532	-0.16	0.871	-5.325133	4.510701
3		-1.824405	1.59875	-1.14	0.255	-4.97126	1.32245
4		-.0221479	1.635092	-0.01	0.989	-3.240536	3.19624
5		-3.712085	1.666385	-2.23	0.027	-6.992068	-.4321023
6		-1.698471	1.945767	-0.87	0.383	-5.528368	2.131426
7		-.3151298	1.584276	-0.20	0.842	-3.433497	2.803237
8		.0125038	1.665041	0.01	0.994	-3.264834	3.289841
9		-.4398757	1.601754	-0.27	0.784	-3.592644	2.712892
10		-1.569803	1.753705	-0.90	0.371	-5.02166	1.882054
11		-.528838	1.649252	-0.32	0.749	-3.775099	2.717423
12		-.4411627	1.685707	-0.26	0.794	-3.759177	2.876852
13		-.9643449	1.69668	-0.57	0.570	-4.303959	2.375269
_cons		8.127861	3.347048	2.43	0.016	1.539791	14.71593

Measures of Fit for regress of activationtime_trans

Log-Lik Intercept Only:	-955.051	Log-Lik Full Model:	-796.399
D(278):	1592.798	LR(43):	317.303
		Prob > LR:	0.000
R2:	0.619	Adjusted R2:	0.561
AIC:	5.151	AIC*n:	1694.798
BIC:	-18.506	BIC':	-68.072

APPENDIX G: SPECIFIC AIM THREE – SITE MODELS PREDICTING CLINICAL TRIAL ACCRUAL

Appendix G-1: Site-Specific Models Predicting Clinical Trial Accrual

Appendix G-1a: Site 104647

```
Negative binomial regression      Number of obs   =      365
                                   LR chi2(26)      =      222.55
Dispersion = mean                 Prob > chi2     =      0.0000
Log likelihood = -1082.8749        Pseudo R2      =      0.0932
```

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

irb						
Local	-.0555245	.1387942	-0.40	0.689	-.3275562	.2165072
sponsor_type						
Institutitonal	-.7644006	.2517369	-3.04	0.002	-1.257796	-.2710053
Industry	-.6812043	.2147786	-3.17	0.002	-1.102163	-.260246
National Group	-.99537	.2554056	-3.90	0.000	-1.495956	-.4947841
placebo						
Yes	-.4884409	.1899457	-2.57	0.010	-.8607275	-.1161542
natenroll_trans	.5510807	.061265	9.00	0.000	.4310034	.671158
natsitescat						
10-49 sites	-.4569062	.15936	-2.87	0.004	-.7692462	-.1445663
50-199 sites	-1.141764	.1986864	-5.75	0.000	-1.531182	-.7523453
200+ sites	-1.689197	.3017838	-5.60	0.000	-2.280683	-1.097712
Unknown	-.6412347	.4551394	-1.41	0.159	-1.533292	.2508222
totalmo_trans						
totalmo_trans	.4322988	.1182335	3.66	0.000	.2005653	.6640323
moaccrdone_trans						
moaccrdone_trans	-.2552333	.0380524	-6.71	0.000	-.3298145	-.180652
primary_purpose						
Supportive Care	.5684731	.326965	1.74	0.082	-.0723665	1.209313
disease_team						
1	.8358012	.3800127	2.20	0.028	.09099	1.580612
2	.0363647	.3324366	0.11	0.913	-.615199	.6879284
3	-.4510802	.3892788	-1.16	0.247	-1.214053	.3118921
4	-.0014715	.423517	-0.00	0.997	-.8315495	.8286066

irb							
Local	.3009099	.1868541	1.61	0.107	-.0653175	.6671373	
sponsor_type							
Instititutional	.2957309	.204066	1.45	0.147	-.1042311	.6956929	
Industry	-.6206198	.1957051	-3.17	0.002	-1.004195	-.237045	
National Group	-.3743732	.206881	-1.81	0.070	-.7798524	.0311061	
friend							
Efficacy	-.153349	.1934602	-0.79	0.428	-.532524	.2258259	
Other	.0391395	.337918	0.12	0.908	-.6231677	.7014467	
natenroll_trans	.614152	.0631274	9.73	0.000	.4904247	.7378794	
natsitescat							
10-49 sites	-.892376	.1820423	-4.90	0.000	-1.249172	-.5355796	
50-199 sites	-1.239667	.2274996	-5.45	0.000	-1.685558	-.7937759	
200+ sites	-1.201846	.2744302	-4.38	0.000	-1.739719	-.6639728	
Unknown	-.495329	.338999	-1.46	0.144	-1.159755	.1690968	
totalmo_trans	.2718281	.1039186	2.62	0.009	.0681513	.4755049	
moaccrdone_trans	-.2837114	.0362657	-7.82	0.000	-.3547909	-.212632	
primary_purpose							
Supportive Care	.9148456	.3358974	2.72	0.006	.2564988	1.573192	
disease_team							
1	-.2472329	.2970308	-0.83	0.405	-.8294025	.3349368	
2	-1.239021	.9510802	-1.30	0.193	-3.103104	.6250623	
3	.1062905	.2816491	0.38	0.706	-.4457317	.6583126	
4	.0379347	.2460064	0.15	0.877	-.4442291	.5200985	
5	-.4725159	.2640246	-1.79	0.074	-.9899946	.0449629	
6	-.8608123	1.302241	-0.66	0.509	-3.413158	1.691533	
7	.1934509	.2572356	0.75	0.452	-.3107217	.6976235	
8	.2121552	.2639944	0.80	0.422	-.3052644	.7295748	
9	-.1018123	.3284605	-0.31	0.757	-.7455831	.5419584	
10	-.9888071	.4799799	-2.06	0.039	-1.92955	-.0480639	
11	-.4175677	.4322202	-0.97	0.334	-1.264704	.4295684	
12	-.7561102	.2803358	-2.70	0.007	-1.305558	-.2066622	
13	.312283	.2734359	1.14	0.253	-.2236416	.8482075	
14	-.3870436	.2772492	-1.40	0.163	-.930442	.1563548	
15	-.320277	.5021592	-0.64	0.524	-1.304491	.6639368	
16	.1335866	.4871137	0.27	0.784	-.8211387	1.088312	
17	-.0578739	.2638335	-0.22	0.826	-.5749781	.4592302	

_cons	-0.4655778	.5740168	-0.81	0.417	-1.59063	.6594744
/lnalpha	-0.4844998	.1000277			-.6805505	-.288449
alpha	.6160053	.0616176			.5063382	.749425

LR test of alpha=0: chibar2(01) = 1216.32 Prob >= chibar2 = 0.000

Measures of Fit for nbreg of local_accrual

Log-Lik Intercept Only:	-1264.475	Log-Lik Full Model:	-1094.047
D(361):	2188.094	LR(37):	340.856
		Prob > LR:	0.000
McFadden's R2:	0.135	McFadden's Adj R2:	0.098
Maximum Likelihood R2:	0.567	Cragg & Uhler's R2:	0.568
AIC:	5.602	AIC*n:	2280.094
BIC:	18.912	BIC':	-118.530

Appendix G-1c: Site 448155

Negative binomial regression	Number of obs	=	697
	LR chi2(38)	=	564.68
Dispersion = mean	Prob > chi2	=	0.0000
Log likelihood = -1998.1847	Pseudo R2	=	0.1238

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
phase					
Phase 1	-.3693882	.4225306	-0.87	0.382	-1.197533 .4587564
Phase 1/2	-.1738514	.4218015	-0.41	0.680	-1.000567 .6528644
Phase 2	-.3246177	.4117792	-0.79	0.431	-1.13169 .4824548
Phase 2/3	-.4483698	.513495	-0.87	0.383	-1.454802 .558062
Phase 3	-.4945648	.4204681	-1.18	0.240	-1.318667 .3295376
Phase 4	-.082885	.7380186	-0.11	0.911	-1.529375 1.363605
None	-.7401331	.4661116	-1.59	0.112	-1.653695 .1734288
sponsor_type					
Instititutional	.4248516	.2313944	1.84	0.066	-.028673 .8783762
Industry	-.4584291	.2242085	-2.04	0.041	-.8978697 -.0189885
National Group	-.015107	.2248079	-0.07	0.946	-.4557224 .4255084
peds					
Yes	-.190139	.1430133	-1.33	0.184	-.47044 .090162

natenroll_trans		.5065872	.0403586	12.55	0.000	.4274857	.5856887
natsitescat							
10-49 sites		-.3256248	.116576	-2.79	0.005	-.5541095	-.0971401
50-199 sites		-.6958279	.1419817	-4.90	0.000	-.9741068	-.4175489
200+ sites		-.8082722	.1844923	-4.38	0.000	-1.169871	-.4466739
Unknown		-.8373126	.2442571	-3.43	0.001	-1.316048	-.3585775
totalmo_trans		.3907095	.0702516	5.56	0.000	.2530188	.5284002
moaccrdone_trans		-.3435651	.0273201	-12.58	0.000	-.3971115	-.2900187
primary_purpose							
Supportive Care		.0081367	.36465	0.02	0.982	-.7065642	.7228377
disease_team							
1		-3.369813	1.214259	-2.78	0.006	-5.749717	-.9899089
2		-.8403173	.1526055	-5.51	0.000	-1.139419	-.541216
3		-.6281789	.5645203	-1.11	0.266	-1.734618	.4782605
4		-.2084333	.1576868	-1.32	0.186	-.5174938	.1006271
5		-.3326303	.1777693	-1.87	0.061	-.6810518	.0157911
6		-.1115767	.2424001	-0.46	0.645	-.5866722	.3635187
7		-.1212197	.1347531	-0.90	0.368	-.3853308	.1428915
8		-.6666943	.1923294	-3.47	0.001	-1.043653	-.2897356
9		-.6310733	.2057029	-3.07	0.002	-1.034244	-.227903
10		.5114419	.3240799	1.58	0.115	-.123743	1.146627
11		.3472078	.225245	1.54	0.123	-.0942643	.7886799
12		-.5054765	.151177	-3.34	0.001	-.801778	-.2091751
13		-2.862236	1.108496	-2.58	0.010	-5.034848	-.6896236
14		-.3358952	.2644559	-1.27	0.204	-.8542193	.1824289
15		-1.533555	.1678319	-9.14	0.000	-1.8625	-1.204611
16		.2155638	.182499	1.18	0.238	-.1421277	.5732554
17		-.3813254	.1546211	-2.47	0.014	-.6843771	-.0782736
18		-.2733534	.6446971	-0.42	0.672	-1.536936	.9902296
19		-.4281826	.1784362	-2.40	0.016	-.777911	-.0784541
_cons		.3155429	.5386075	0.59	0.558	-.7401084	1.371194
/lnalpha		-.5601355	.0746754			-.7064967	-.4137743
alpha		.5711317	.0426495			.4933696	.6611501

LR test of alpha=0: chibar2(01) = 1752.52

Prob >= chibar2 = 0.000

Measures of Fit for nbreg of local_accrual

Log-Lik Intercept Only:	-2280.524	Log-Lik Full Model:	-1998.185
D(651):	3996.369	LR(38):	564.680
		Prob > LR:	0.000
McFadden's R2:	0.124	McFadden's Adj R2:	0.104
Maximum Likelihood R2:	0.555	Cragg & Uhler's R2:	0.556
AIC:	5.866	AIC*n:	4088.369
BIC:	-265.588	BIC':	-315.902

Appendix G-1d: Site 494048

Negative binomial regression	Number of obs	=	552
	LR chi2(36)	=	348.77
Dispersion = mean	Prob > chi2	=	0.0000
Log likelihood = -1455.1689	Pseudo R2	=	0.1070

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

phase						
Phase 1	1.000691	1.032166	0.97	0.332	-1.022317	3.023699
Phase 1/2	1.221762	1.039333	1.18	0.240	-.8152934	3.258817
Phase 2	.776221	1.02945	0.75	0.451	-1.241464	2.793906
Phase 2/3	.6092842	1.143164	0.53	0.594	-1.631277	2.849845
Phase 3	.6122638	1.040742	0.59	0.556	-1.427554	2.652081
Phase 4	1.4139	1.134706	1.25	0.213	-.8100824	3.637883
None	.947351	1.05767	0.90	0.370	-1.125643	3.020345

sponsor_type						
Institutitonal	.3009689	.3381593	0.89	0.373	-.3618111	.9637489
Industry	-.6204869	.3264933	-1.90	0.057	-1.260402	.0194282
National Group	-.8231966	.3380058	-2.44	0.015	-1.485676	-.1607175

natenroll_trans	.6114453	.0607751	10.06	0.000	.4923282	.7305623

natsitescat						
10-49 sites	-.6858121	.1542933	-4.44	0.000	-.9882214	-.3834027
50-199 sites	-1.111765	.1986885	-5.60	0.000	-1.501187	-.7223428
200+ sites	-1.168979	.2669461	-4.38	0.000	-1.692184	-.6457746
Unknown	-.5681781	.4023271	-1.41	0.158	-1.356725	.2203686

totalmo_trans	.2782846	.0962541	2.89	0.004	.08963	.4669392
moaccrdone_trans	-.2441671	.0382413	-6.38	0.000	-.3191188	-.1692155

primary_purpose							
Supportive Care		.0886171	.3083742	0.29	0.774	-.5157852	.6930195
disease_team							
1		-2.11719	.6071901	-3.49	0.000	-3.307261	-.9271197
2		-1.817452	.5745144	-3.16	0.002	-2.943479	-.6914244
4		-2.667098	.6122274	-4.36	0.000	-3.867042	-1.467155
5		-1.545001	.6121894	-2.52	0.012	-2.74487	-.3451314
6		-1.611253	1.240317	-1.30	0.194	-4.04223	.8197248
7		-1.575201	.6017739	-2.62	0.009	-2.754657	-.3957461
8		-1.13868	.6671145	-1.71	0.088	-2.4462	.1688404
9		-1.117873	.6204588	-1.80	0.072	-2.333949	.0982042
10		-2.349178	.6098908	-3.85	0.000	-3.544542	-1.153814
11		-1.858414	.7024345	-2.65	0.008	-3.23516	-.4816677
12		-1.918662	.5996176	-3.20	0.001	-3.09389	-.7434327
13		-3.380944	.741754	-4.56	0.000	-4.834755	-1.927133
14		-1.964342	.6045185	-3.25	0.001	-3.149177	-.7795077
15		-1.765049	.6359878	-2.78	0.006	-3.011562	-.5185362
16		-2.543334	.8014867	-3.17	0.002	-4.114219	-.9724487
17		-1.933938	.9896588	-1.95	0.051	-3.873633	.0057579
18		-1.813025	.6905789	-2.63	0.009	-3.166534	-.459515
19		-1.786249	.6077201	-2.94	0.003	-2.977358	-.595139
_cons		.4637986	1.237771	0.37	0.708	-1.962187	2.889785
/lnalpha		-.1447455	.0829887			-.3074004	.0179093
alpha		.8652424	.0718053			.7353561	1.018071

LR test of alpha=0: chibar2(01) = 1679.63 Prob >= chibar2 = 0.000

Measures of Fit for nbreg of local_accrual

Log-Lik Intercept Only:	-1629.555	Log-Lik Full Model:	-1455.169
D(509):	2910.338	LR(36):	348.772
		Prob > LR:	0.000
McFadden's R2:	0.107	McFadden's Adj R2:	0.081
Maximum Likelihood R2:	0.468	Cragg & Uhler's R2:	0.470
AIC:	5.428	AIC*n:	2996.338
BIC:	-303.258	BIC':	-121.484

Appendix G-1e: Site 512786

Negative binomial regression	Number of obs	=	484
	LR chi2(33)	=	293.91
Dispersion = mean	Prob > chi2	=	0.0000
Log likelihood = -1576.2356	Pseudo R2	=	0.0853

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
phase						
Phase 1	.0238747	.2840517	0.08	0.933	-.5328564	.5806059
Phase 1/2	.3731285	.2827624	1.32	0.187	-.1810757	.9273327
Phase 2	.0635337	.2590053	0.25	0.806	-.4441073	.5711747
Phase 2/3	.1592321	.7262724	0.22	0.826	-1.264236	1.5827
Phase 3	.1839785	.3121293	0.59	0.556	-.4277837	.7957408
Phase 4	.8910371	.5800379	1.54	0.124	-.2458162	2.02789
irb						
Local	.23849	.1034697	2.30	0.021	.0356931	.4412869
sponsor_type						
Institutional	.2715557	.1822501	1.49	0.136	-.085648	.6287593
Industry	-.460929	.1689898	-2.73	0.006	-.792143	-.129715
National Group	-.4535349	.2263842	-2.00	0.045	-.8972398	-.00983
natenroll_trans	.2523652	.0536664	4.70	0.000	.147181	.3575494
natsitescat						
10-49 sites	-.2679925	.1312281	-2.04	0.041	-.5251948	-.0107902
50-199 sites	-.3560904	.1967585	-1.81	0.070	-.74173	.0295491
200+ sites	-.3923535	.2959884	-1.33	0.185	-.9724801	.1877732
Unknown	-.0835072	.2024471	-0.41	0.680	-.4802962	.3132819
totalmo_trans	.5480934	.0843056	6.50	0.000	.3828575	.7133293
moaccrdone_trans	-.2503641	.029485	-8.49	0.000	-.3081538	-.1925745
disease_team						
2	-.1067287	.2611665	-0.41	0.683	-.6186056	.4051483
3	.652127	.1965326	3.32	0.001	.2669301	1.037324
4	.2081555	.20856	1.00	0.318	-.2006146	.6169257
5	-.253699	.246602	-1.03	0.304	-.7370301	.2296321
6	-.2396966	.2979853	-0.80	0.421	-.823737	.3443438
7	.0064793	.3946523	0.02	0.987	-.7670251	.7799837
8	.4643643	.9742177	0.48	0.634	-1.445067	2.373796

9		.6170586	.5969222	1.03	0.301	-.5528874	1.787005
10		.0266904	.1824248	0.15	0.884	-.3308557	.3842365
11		-.5697493	1.02241	-0.56	0.577	-2.573637	1.434138
12		-.1118027	.2382793	-0.47	0.639	-.5788216	.3552163
13		.4175434	.2137141	1.95	0.051	-.0013285	.8364152
14		-.6858171	.2723293	-2.52	0.012	-1.219573	-.1520615
15		.0626144	.2883424	0.22	0.828	-.5025263	.627755
16		-1.271912	.6916317	-1.84	0.066	-2.627485	.0836614
17		.0765005	.2067984	0.37	0.711	-.3288169	.4818179
_cons		-.2482	.4468813	-0.56	0.579	-1.124071	.6276713

/lnalpha		-.3095109	.07696			-.4603498	-.158672

alpha		.7338058	.0564737			.6310629	.8532762

LR test of alpha=0: chibar2(01) = 2521.34 Prob >= chibar2 = 0.000

Measures of Fit for nbreg of local_accrual

Log-Lik Intercept Only:	-1723.190	Log-Lik Full Model:	-1576.236
D(444):	3152.471	LR(33):	293.909
		Prob > LR:	0.000
McFadden's R2:	0.085	McFadden's Adj R2:	0.062
Maximum Likelihood R2:	0.455	Cragg & Uhler's R2:	0.456
AIC:	6.679	AIC*n:	3232.471
BIC:	407.625	BIC':	-89.900

Appendix G-1f: Site 560623

Negative binomial regression		Number of obs	=	179
		LR chi2(28)	=	111.69
Dispersion = mean		Prob > chi2	=	0.0000
Log likelihood = -276.05464		Pseudo R2	=	0.1683

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
phase						
Phase 1/2	-.1890346	.5044713	-0.37	0.708	-1.17778	.799711
Phase 2	-.2425543	.573109	-0.42	0.672	-1.365827	.8807188
Phase 2/3	-.4431795	.9384069	-0.47	0.637	-2.282423	1.396064
Phase 3	.0240299	.5973593	0.04	0.968	-1.146773	1.194833
Phase 4	-21.01789	45061.49	-0.00	1.000	-88339.92	88297.88
None	-.1688369	.9224158	-0.18	0.855	-1.976739	1.639065

sponsor_type							
Industry		-.4048534	.5223806	-0.78	0.438	-1.428701	.6189939
National Group		-.4519362	.5273485	-0.86	0.391	-1.48552	.5816479

priend							
Efficacy		.0103822	.6077847	0.02	0.986	-1.180854	1.201618
Other		1.588991	1.611245	0.99	0.324	-1.568992	4.746974

natenroll_trans		.5940102	.1500874	3.96	0.000	.2998443	.8881761

natsitescat							
10-49 sites		-.3138507	.434376	-0.72	0.470	-1.165212	.5375106
50-199 sites		-1.084828	.4972709	-2.18	0.029	-2.059461	-.1101945
200+ sites		-.9379467	.5521854	-1.70	0.089	-2.02021	.1443167
Unknown		.2446314	.9606545	0.25	0.799	-1.638217	2.12748

totalmo_trans		.4501827	.2714583	1.66	0.097	-.0818657	.9822312
moaccrdone_trans		-.3022845	.0708456	-4.27	0.000	-.4411393	-.1634297

primary_purpose							
Supportive Care		-.8980197	.6477061	-1.39	0.166	-2.1675	.3714609

disease_team							
1		1.586854	.4555832	3.48	0.000	.6939271	2.47978
2		1.287789	.5698801	2.26	0.024	.1708449	2.404734
3		1.273404	.4526145	2.81	0.005	.3862955	2.160512
4		.8899164	.4170041	2.13	0.033	.0726033	1.707229
5		2.815293	.5125071	5.49	0.000	1.810798	3.819789
6		.4042761	.5610097	0.72	0.471	-.6952828	1.503835
7		1.290976	.4478104	2.88	0.004	.4132841	2.168669
8		1.629798	1.243558	1.31	0.190	-.8075312	4.067127
9		2.290214	.7772869	2.95	0.003	.7667602	3.813669
10		-20.43243	45061.49	-0.00	1.000	-88339.33	88298.47

_cons		-3.616622	1.449076	-2.50	0.013	-6.456758	-.7764862

/lnalpha		-.4441159	.2330047			-.9007967	.0125648

alpha		.6413911	.1494471			.4062459	1.012644

LR test of alpha=0: chibar2(01) = 78.79				Prob >= chibar2 = 0.000			

Measures of Fit for nbreg of local_accrual

Log-Lik Intercept Only:	-331.899	Log-Lik Full Model:	-276.055
D(143):	552.109	LR(28):	111.689
		Prob > LR:	0.000
McFadden's R2:	0.168	McFadden's Adj R2:	0.060
Maximum Likelihood R2:	0.464	Cragg & Uhler's R2:	0.476
AIC:	3.487	AIC*n:	624.109
BIC:	-189.687	BIC':	33.558

Appendix G-1g: Site 575415

Negative binomial regression	Number of obs	=	410
	LR chi2(28)	=	296.70
Dispersion = mean	Prob > chi2	=	0.0000
Log likelihood = -1131.828	Pseudo R2	=	0.1159

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

irb						
Local	.1059458	.1906911	0.56	0.578	-.2678018	.4796934
sponsor_type						
Institutitonal	-.2421846	.2271986	-1.07	0.286	-.6874857	.2031165
Industry	-.9896439	.2635384	-3.76	0.000	-1.50617	-.4731182
National Group	-1.230769	.2654572	-4.64	0.000	-1.751056	-.7104826
peds						
Yes	-.0202699	.2112847	-0.10	0.924	-.4343803	.3938406
random						
Yes	-.1978224	.14844	-1.33	0.183	-.4887595	.0931147
natenroll_trans	.551405	.0559803	9.85	0.000	.4416856	.6611244
natsitescat						
10-49 sites	-.2671798	.1663133	-1.61	0.108	-.5931479	.0587883
50-199 sites	-.8519705	.2174587	-3.92	0.000	-1.278182	-.4257593
200+ sites	-1.389022	.2787604	-4.98	0.000	-1.935382	-.8426616
Unknown	.0459851	.2781557	0.17	0.869	-.4991902	.5911603
totalmo_trans	.3037763	.1040091	2.92	0.003	.0999221	.5076304
moaccrdone_trans	-.2937358	.0390519	-7.52	0.000	-.3702762	-.2171954
disease_team						

1		.0841342	.6726128	0.13	0.900	-1.234163	1.402431
2		.1066082	.281654	0.38	0.705	-.4454236	.65864
3		-1.131312	.7771079	-1.46	0.145	-2.654416	.3917916
4		.4245582	.2540908	1.67	0.095	-.0734507	.922567
5		-.0368377	.2429565	-0.15	0.879	-.5130238	.4393483
6		.2853069	.2252698	1.27	0.205	-.1562137	.7268275
7		.2317329	.2767597	0.84	0.402	-.3107062	.774172
8		-.0518323	.2270609	-0.23	0.819	-.4968636	.3931989
9		.418876	.257459	1.63	0.104	-.0857344	.9234864
10		1.324823	.9251299	1.43	0.152	-.4883986	3.138044
11		-.2596721	.4837785	-0.54	0.591	-1.207861	.6885163
12		-.5784005	.2633083	-2.20	0.028	-1.094475	-.0623256
13		-.4136314	.318544	-1.30	0.194	-1.037966	.2107034
14		-.7867787	.373308	-2.11	0.035	-1.518449	-.0551086
15		-.2698436	.4431613	-0.61	0.543	-1.138424	.5987365
_cons		.140829	.4974039	0.28	0.777	-.8340647	1.115723

/lnalpha		-.2880243	.0969527			-.478048	-.0980006

alpha		.7497434	.0726896			.6199924	.9066484

LR test of alpha=0: chibar2(01) = 1184.60 Prob >= chibar2 = 0.000

Measures of Fit for nbreg of local_accrual

Log-Lik Intercept Only:	-1280.179	Log-Lik Full Model:	-1131.828
D(374):	2263.656	LR(28):	296.701
		Prob > LR:	0.000
McFadden's R2:	0.116	McFadden's Adj R2:	0.088
Maximum Likelihood R2:	0.515	Cragg & Uhler's R2:	0.516
AIC:	5.697	AIC*n:	2335.656
BIC:	13.613	BIC':	-128.249

Appendix G-1h: Site 598430

Negative binomial regression		Number of obs	=	326
		LR chi2(33)	=	158.63
Dispersion = mean		Prob > chi2	=	0.0000
Log likelihood = -721.37973		Pseudo R2	=	0.0991

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]

phase					

Phase 1/2	.3514323	.4487117	0.78	0.434	-.5280265	1.230891
Phase 2	-.196397	.2874345	-0.68	0.494	-.7597583	.3669642
Phase 2/3	-.676317	.7603173	-0.89	0.374	-2.166512	.8138775
Phase 3	-.3448489	.3541914	-0.97	0.330	-1.039051	.3493534
Phase 4	-.4935263	.8139045	-0.61	0.544	-2.08875	1.101697
None	.5383356	.5682294	0.95	0.343	-.5753736	1.652045
irb						
Local	.1779674	.4085574	0.44	0.663	-.6227904	.9787252
sponsor_type						
Institutional	.1828585	.3893728	0.47	0.639	-.5802982	.9460153
Industry	-.9168694	.5329658	-1.72	0.085	-1.961463	.1277244
National Group	-.9355007	.380615	-2.46	0.014	-1.681493	-.189509
peds						
Yes	.3925437	.2823849	1.39	0.164	-.1609206	.9460079
natenroll_trans	.4875717	.0903525	5.40	0.000	.310484	.6646593
natsitescat						
10-49 sites	-.8697358	.3308729	-2.63	0.009	-1.518235	-.2212368
50-199 sites	-.96423	.3716573	-2.59	0.009	-1.692665	-.2357951
200+ sites	-.858403	.4263195	-2.01	0.044	-1.693974	-.0228323
Unknown	-.4143899	.4888598	-0.85	0.397	-1.372538	.5437577
totalmo_trans	.2066294	.1421503	1.45	0.146	-.07198	.4852388
moaccrdone_trans	-.2367306	.0473423	-5.00	0.000	-.3295198	-.1439414
primary_purpose						
Supportive Care	-.3301417	.5165948	-0.64	0.523	-1.342649	.6823656
disease_team						
1	-.4209589	.4215382	-1.00	0.318	-1.247159	.4052408
2	-.5429432	.395331	-1.37	0.170	-1.317778	.2318914
3	-1.229533	.5437084	-2.26	0.024	-2.295182	-.1638842
4	-.1742623	.3989662	-0.44	0.662	-.9562216	.6076971
5	-.1597124	.3676284	-0.43	0.664	-.8802509	.5608261
6	.0141927	.3997887	0.04	0.972	-.7693789	.7977642
7	-.4289432	.4226842	-1.01	0.310	-1.257389	.3995026
8	.4491728	.4574532	0.98	0.326	-.4474189	1.345765
9	-.1236466	.5106229	-0.24	0.809	-1.124449	.8771559
10	-.1014738	.5611468	-0.18	0.856	-1.201301	.9983537
11	.9572945	1.096947	0.87	0.383	-1.192682	3.107271
12	.3845161	.4478785	0.86	0.391	-.4933096	1.262342
13	-.9616073	.4679372	-2.05	0.040	-1.878747	-.0444671

14		-.4915948	.3967475	-1.24	0.215	-1.269206	.286016
_cons		.3292436	.7852966	0.42	0.675	-1.209909	1.868397
/lnalpha		-.1348601	.1274126			-.3845843	.114864
alpha		.8738381	.111338			.6807336	1.121721

LR test of alpha=0: chibar2(01) = 383.78 Prob >= chibar2 = 0.000

Measures of Fit for nbreg of local_accrual

Log-Lik Intercept Only:	-800.695	Log-Lik Full Model:	-721.380
D(284):	1442.759	LR(33):	158.630
		Prob > LR:	0.000
McFadden's R2:	0.099	McFadden's Adj R2:	0.047
Maximum Likelihood R2:	0.385	Cragg & Uhler's R2:	0.388
AIC:	4.683	AIC*n:	1526.759
BIC:	-200.719	BIC':	32.337

Appendix G-1i: Site 602591

Negative binomial regression	Number of obs	=	395
	LR chi2(23)	=	254.65
Dispersion = mean	Prob > chi2	=	0.0000
Log likelihood = -950.6596	Pseudo R2	=	0.1181

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

sponsor_type						
Institutitonal	-.413279	.2466561	-1.68	0.094	-.896716 .070158	
Industry	-1.351568	.2494122	-5.42	0.000	-1.840407 -.8627291	
National Group	-1.248505	.2670499	-4.68	0.000	-1.771913 -.7250963	
natenroll_trans	.5486444	.0566469	9.69	0.000	.4376186 .6596703	
natsitescat						
10-49 sites	-.3035362	.1640237	-1.85	0.064	-.6250167 .0179443	
50-199 sites	-.7889316	.19737	-4.00	0.000	-1.17577 -.4020934	
200+ sites	-1.097642	.2669301	-4.11	0.000	-1.620815 -.5744685	
Unknown	-.667696	.3443197	-1.94	0.052	-1.34255 .0071581	
totalmo_trans	.2609851	.1108565	2.35	0.019	.0437104 .4782598	
moaccrdone_trans	-.4045044	.0393328	-10.28	0.000	-.4815953 -.3274134	

Phase 1	-.3911047	.7578812	-0.52	0.606	-1.876525	1.094315
Phase 1/2	-.1092503	.7535295	-0.14	0.885	-1.586141	1.36764
Phase 2	-.3626771	.7516537	-0.48	0.629	-1.835891	1.110537
Phase 2/3	-.5857072	1.06466	-0.55	0.582	-2.672403	1.500989
Phase 3	-.8193224	.766509	-1.07	0.285	-2.321652	.6830076
Phase 4	.109982	1.097141	0.10	0.920	-2.040375	2.260339
None	-.1814073	.7894878	-0.23	0.818	-1.728775	1.36596
sponsor_type						
Instititutional	-.3759014	.1833995	-2.05	0.040	-.7353578	-.016445
Industry	-.6043809	.1294084	-4.67	0.000	-.8580168	-.350745
National Group	-.5177852	.1615446	-3.21	0.001	-.8344067	-.2011637
peds						
Yes	.0037365	.2070513	0.02	0.986	-.4020765	.4095496
random						
Yes	.0111475	.1074347	0.10	0.917	-.1994207	.2217157
natenroll_trans	.5339427	.0477495	11.18	0.000	.4403554	.6275301
natsitescat						
10-49 sites	-.5683534	.1248684	-4.55	0.000	-.813091	-.3236157
50-199 sites	-.7657871	.1558689	-4.91	0.000	-1.071285	-.4602897
200+ sites	-.8829576	.2052644	-4.30	0.000	-1.285269	-.4806467
Unknown	-.8624442	.2541691	-3.39	0.001	-1.360606	-.364282
totalmo_trans	.3773527	.0785973	4.80	0.000	.2233048	.5314005
moaccrdone_trans	-.3604784	.0304869	-11.82	0.000	-.4202317	-.3007251
disease_team						
1	-.2496689	.2540652	-0.98	0.326	-.7476275	.2482897
2	-.2474498	.2328592	-1.06	0.288	-.7038455	.2089458
3	-.2608517	.3616558	-0.72	0.471	-.969684	.4479805
4	-.0576509	.1644694	-0.35	0.726	-.380005	.2647033
5	-.3689491	.192574	-1.92	0.055	-.7463872	.008489
6	-.3068241	.2049464	-1.50	0.134	-.7085116	.0948634
7	-.124864	.3239251	-0.39	0.700	-.7597454	.5100175
8	-.5921995	.1906074	-3.11	0.002	-.9657831	-.2186159
9	-.3353258	.1965723	-1.71	0.088	-.7206004	.0499488
10	.0302682	.1491596	0.20	0.839	-.2620791	.3226156
11	.1118257	.2361547	0.47	0.636	-.3510291	.5746804
12	-1.313846	.2451252	-5.36	0.000	-1.794282	-.8334093
13	.0184282	.202302	0.09	0.927	-.3780764	.4149327
14	.144251	.2694147	0.54	0.592	-.383792	.672294
15	-.4861042	.22133	-2.20	0.028	-.919903	-.0523054

_cons		.8612445	.8053465	1.07	0.285	-.7172057	2.439695
/lnalpha		-.7950652	.079465			-.9508137	-.6393168
alpha		.4515518	.0358825			.3864265	.5276528

LR test of alpha=0: chibar2(01) = 1981.81 Prob >= chibar2 = 0.000

Measures of Fit for nbreg of local_accrual

Log-Lik Intercept Only:	-1693.479	Log-Lik Full Model:	-1473.771
D(414):	2947.543	LR(34):	439.414
		Prob > LR:	0.000
McFadden's R2:	0.130	McFadden's Adj R2:	0.105
Maximum Likelihood R2:	0.618	Cragg & Uhler's R2:	0.619
AIC:	6.648	AIC*n:	3031.543
BIC:	412.831	BIC':	-231.250

Appendix G-1k: Site 696337

Negative binomial regression	Number of obs	=	182
	LR chi2(22)	=	90.63
Dispersion = mean	Prob > chi2	=	0.0000
Log likelihood = -429.92362	Pseudo R2	=	0.0954

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

sponsor_type						
Institutitonal	-2.016601	.9464522	-2.13	0.033	-3.871613	-.1615885
Industry	-.5489011	.5901912	-0.93	0.352	-1.705654	.6078524
National Group	-.7337731	.6316352	-1.16	0.245	-1.971755	.5042093
peds						
Yes	.0468469	.4586068	0.10	0.919	-.8520059	.9456997
random						
Yes	-.3798285	.2197411	-1.73	0.084	-.8105131	.0508561
friend						
Efficacy	.0265886	.3321871	0.08	0.936	-.6244862	.6776634
Other	.8569333	.5187457	1.65	0.099	-.1597896	1.873656
natenroll_trans	.5496809	.107206	5.13	0.000	.3395609	.7598009

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
phase						
Phase 1/2	.2409437	.1908549	1.26	0.207	-.133125	.6150125
Phase 2	.0318709	.14214	0.22	0.823	-.2467183	.3104601
Phase 2/3	-.0678859	.3433899	-0.20	0.843	-.7409178	.605146
Phase 3	-.3192724	.193639	-1.65	0.099	-.6987979	.0602531
Phase 4	-.0272158	.4738343	-0.06	0.954	-.9559139	.9014823
None	.1654504	.2435628	0.68	0.497	-.3119239	.6428246
sponsor_type						
Instititutional	.0518868	.190048	0.27	0.785	-.3206005	.4243741
Industry	-.8020536	.2016129	-3.98	0.000	-1.197208	-.4068996
National Group	-.8222982	.2242019	-3.67	0.000	-1.261726	-.3828707
placebo						
Yes	-.4813167	.1606784	-3.00	0.003	-.7962406	-.1663928
natenroll_trans	.6961303	.0465178	14.96	0.000	.6049571	.7873035
natsitescat						
10-49 sites	-.8281526	.1484754	-5.58	0.000	-1.119159	-.5371462
50-199 sites	-1.073653	.1838211	-5.84	0.000	-1.433936	-.7133705
200+ sites	-1.658757	.228593	-7.26	0.000	-2.106792	-1.210723
Unknown	-.9338967	.3710417	-2.52	0.012	-1.661125	-.2066682
totalmo_trans	.2009115	.0768463	2.61	0.009	.0502956	.3515274
moaccrdone_trans	-.3287862	.0304635	-10.79	0.000	-.3884935	-.2690788
primary_purpose						
Supportive Care	-.089499	.2540407	-0.35	0.725	-.5874097	.4084117
_cons	.1105425	.3380876	0.33	0.744	-.552097	.7731821
/lnalpha	-.4025646	.0826673			-.5645895	-.2405396
alpha	.6686032	.0552716			.5685935	.7862035

LR test of alpha=0: chibar2(01) = 3054.91 Prob >= chibar2 = 0.000

Measures of Fit for nbreg of local_accrual

Log-Lik Intercept Only:	-1741.380	Log-Lik Full Model:	-1511.054
D(504):	3022.108	LR(18):	460.652
		Prob > LR:	0.000
McFadden's R2:	0.132	McFadden's Adj R2:	0.118
Maximum Likelihood R2:	0.581	Cragg & Uhler's R2:	0.582
AIC:	5.807	AIC*n:	3072.108
BIC:	-138.470	BIC':	-347.774

Appendix G-1m: Site 715532

Negative binomial regression	Number of obs	=	225
	LR chi2(21)	=	187.26
Dispersion = mean	Prob > chi2	=	0.0000
Log likelihood = -511.66891	Pseudo R2	=	0.1547

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

sponsor_type						
Institutional	1.624316	1.778418	0.91	0.361	-1.861318 5.109951	
Industry	3.534986	1.275183	2.77	0.006	1.035673 6.0343	
National Group	.8175742	1.29508	0.63	0.528	-1.720736 3.355884	
peds						
Yes	-.5194587	.302935	-1.71	0.086	-1.1132 .0742829	
natenroll_trans	1.171886	.3185365	3.68	0.000	.5475655 1.796206	
natsitescat						
10-49 sites	-.8697844	.3582539	-2.43	0.015	-1.571949 -.1676196	
50-199 sites	-.6171174	.3950221	-1.56	0.118	-1.391346 .1571118	
200+ sites	-.4751006	.4287851	-1.11	0.268	-1.315504 .3653028	
Unknown	-1.395895	.628901	-2.22	0.026	-2.628518 -.1632713	
totalmo_trans	.582035	.1791125	3.25	0.001	.2309809 .9330891	
moaccrdone_trans	-.3269586	.0486349	-6.72	0.000	-.4222813 -.2316358	
disease_team						
1	-.1587715	1.436888	-0.11	0.912	-2.97502 2.657477	
2	-1.223813	2.107015	-0.58	0.561	-5.353487 2.905861	
3	-19.05688	16343.54	-0.00	0.999	-32051.8 32013.69	
4	-.8309123	1.693306	-0.49	0.624	-4.149731 2.487906	
5	.6015063	1.611154	0.37	0.709	-2.556297 3.75931	
6	-.5760153	1.634347	-0.35	0.725	-3.779277 2.627246	

7		-.2173445	1.654628	-0.13	0.895	-3.460356	3.025667
natenroll_trans		0	(omitted)				
sponsor_type#c.natenroll_trans							
Instititutional		-.6218661	.4855484	-1.28	0.200	-1.573523	.3297912
Industry		-1.225524	.3470854	-3.53	0.000	-1.905799	-.545249
National Group		-.6581477	.3317241	-1.98	0.047	-1.308315	-.0079803
_cons		-3.448406	1.972446	-1.75	0.080	-7.314328	.4175167

/lnalpha		-.3049672	.1489927			-.5969875	-.0129468

alpha		.7371476	.1098296			.5504674	.9871367

LR test of alpha=0: chibar2(01) = 401.22 Prob >= chibar2 = 0.000

Measures of Fit for nbreg of local_accrual

Log-Lik Intercept Only:	-605.297	Log-Lik Full Model:	-511.669
D(196):	1023.338	LR(21):	187.256
		Prob > LR:	0.000
McFadden's R2:	0.155	McFadden's Adj R2:	0.107
Maximum Likelihood R2:	0.565	Cragg & Uhler's R2:	0.568
AIC:	4.806	AIC*n:	1081.338
BIC:	-38.218	BIC':	-73.518

Appendix G-In: Site 846594

Negative binomial regression	Number of obs	=	294
	LR chi2(27)	=	188.59
Dispersion = mean	Prob > chi2	=	0.0000
Log likelihood = -740.16802	Pseudo R2	=	0.1130

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
sponsor_type						
Instititutional	.2382805	.3224614	0.74	0.460	-.3937323	.8702933
Industry	-.7789488	.3344337	-2.33	0.020	-1.434427	-.1234708
National Group	-.8725146	.3472914	-2.51	0.012	-1.553193	-.1918361
peds						
Yes	.1010909	.1872939	0.54	0.589	-.2659984	.4681802
natenroll_trans	.4503979	.063282	7.12	0.000	.3263675	.5744284

natsitescat							
10-49 sites	-.697449	.2752641	-2.53	0.011	-1.236957	-.1579412	
50-199 sites	-1.22571	.3322285	-3.69	0.000	-1.876866	-.5745545	
200+ sites	-1.529256	.4153365	-3.68	0.000	-2.343301	-.7152117	
Unknown	-.6446439	.4287532	-1.50	0.133	-1.484985	.1956969	
totalmo_trans	.4221316	.1184511	3.56	0.000	.1899717	.6542915	
moaccrdone_trans	-.141527	.0466902	-3.03	0.002	-.2330381	-.050016	
primary_purpose							
Supportive Care	.135873	.3672858	0.37	0.711	-.5839939	.85574	
precision							
Yes	-16.83654	2394.526	-0.01	0.994	-4710.021	4676.348	
disease_team							
1	-.8084183	.2282351	-3.54	0.000	-1.255751	-.3610858	
2	-.642779	.2672269	-2.41	0.016	-1.166534	-.119024	
3	-.7950768	.2985031	-2.66	0.008	-1.380132	-.2100216	
4	-.4177502	.3052815	-1.37	0.171	-1.016091	.1805906	
5	-.647868	.3157971	-2.05	0.040	-1.266819	-.028917	
6	.063935	.2703476	0.24	0.813	-.4659364	.5938065	
7	1.3301	.4473125	2.97	0.003	.4533835	2.206816	
8	.4024023	.6434856	0.63	0.532	-.8588063	1.663611	
9	-.5730694	.4310443	-1.33	0.184	-1.417901	.2717618	
10	-.0575509	.6983216	-0.08	0.934	-1.426236	1.311134	
11	-.6886204	.8114746	-0.85	0.396	-2.279081	.9018406	
12	-.9132617	.2891294	-3.16	0.002	-1.479945	-.3465785	
13	-1.039438	.9962191	-1.04	0.297	-2.991992	.9131155	
14	-.4890855	.3097965	-1.58	0.114	-1.096275	.1181045	
_cons	-.0984721	.573738	-0.17	0.864	-1.222978	1.026034	

/lnalpha	-.1736924	.1120665			-.3933387	.0459538	

alpha	.8405554	.0941981			.6748002	1.047026	

LR test of alpha=0: chibar2(01) = 875.05					Prob >= chibar2 = 0.000		

Measures of Fit for nbreg of local_accrual

Log-Lik Intercept Only:	-290.003	Log-Lik Full Model:	-254.263
D(77):	508.526	LR(11):	71.480
		Prob > LR:	0.000
McFadden's R2:	0.123	McFadden's Adj R2:	0.068
Maximum Likelihood R2:	0.536	Cragg & Uhler's R2:	0.537
AIC:	5.812	AIC*n:	540.526
BIC:	159.516	BIC':	-21.621

Appendix G-1p: Site 998666

Negative binomial regression		Number of obs	=	193
		LR chi2(21)	=	121.02
Dispersion = mean		Prob > chi2	=	0.0000
Log likelihood = -463.3665		Pseudo R2	=	0.1155

local_accrual	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

irb						
Local	.2510573	.2061431	1.22	0.223	-.1529758	.6550904
sponsor_type						
Instititutional	-1.030795	.5941793	-1.73	0.083	-2.195365	.1337746
Industry	-2.079804	.4376827	-4.75	0.000	-2.937646	-1.221961
National Group	-1.53639	.4574262	-3.36	0.001	-2.432929	-.639851
natenroll_trans						
	.4778684	.083344	5.73	0.000	.3145171	.6412198
natsitescat						
10-49 sites	-.5585292	.2828636	-1.97	0.048	-1.112932	-.0041267
50-199 sites	-1.238561	.332162	-3.73	0.000	-1.889587	-.5875358
200+ sites	-1.651961	.4272422	-3.87	0.000	-2.48934	-.8145819
Unknown	-1.075999	.4119875	-2.61	0.009	-1.88348	-.2685184
totalmo_trans						
	.4095067	.1848277	2.22	0.027	.0472511	.7717623
moaccrdone_trans						
	-.2284901	.0672859	-3.40	0.001	-.3603679	-.0966122
disease_team						
1	-.69288	.3797354	-1.82	0.068	-1.437148	.0513877
2	-.417919	.3416759	-1.22	0.221	-1.087591	.2517534
3	-.550553	.3734397	-1.47	0.140	-1.282481	.1813753
4	-1.223076	.4391554	-2.79	0.005	-2.083804	-.3623468

5		-.7123758	.3795795	-1.88	0.061	-1.456338	.0315863
6		.1806929	.3871327	0.47	0.641	-.5780732	.939459
7		-1.684348	.573329	-2.94	0.003	-2.808052	-.5606433
8		.4723203	.5101232	0.93	0.355	-.5275027	1.472143
9		-.3520085	.4139512	-0.85	0.395	-1.163338	.4593211
10		.5403136	.3498181	1.54	0.122	-.1453174	1.225945
_cons		.9444198	.7127793	1.32	0.185	-.4526019	2.341442

/lnalpha		-.2074955	.1612612			-.5235616	.1085706

alpha		.8126169	.1310436			.5924068	1.114684

LR test of alpha=0: chibar2(01) = 288.40 Prob >= chibar2 = 0.000

Measures of Fit for nbreg of local_accrual

Log-Lik Intercept Only:	-523.878	Log-Lik Full Model:	-463.366
D(166):	926.733	LR(21):	121.023
		Prob > LR:	0.000
McFadden's R2:	0.116	McFadden's Adj R2:	0.064
Maximum Likelihood R2:	0.466	Cragg & Uhler's R2:	0.468
AIC:	5.082	AIC*n:	980.733
BIC:	53.126	BIC':	-10.507

Appendix G-2: Site-Specific Prediction Model versus Actual Accrual

Appendix G-2a: Site 104647

Actual Accrual	Site Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	1	21	11	4	7	2	1	47
1 < 4	2	47	30	13	12	1	0	105
4 < 7	0	16	27	18	8	1	0	70
7 < 10	0	9	15	4	7	2	0	37
10 < 20	0	4	16	12	26	3	0	61
20 < 50	0	0	4	0	20	12	0	36
> 50	0	0	0	0	1	4	4	9
Total	3	97	103	51	81	25	5	365

Appendix G-2b: Site 173472

Actual Accrual	Site Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	13	42	11	1	2	1	1	71
1 < 4	12	75	30	6	9	3	0	135
4 < 7	0	22	23	10	4	1	0	60
7 < 10	0	10	14	7	10	2	0	43
10 < 20	0	4	13	9	23	4	0	53
20 < 50	0	0	2	2	11	18	3	36
> 50	0	0	0	0	1	4	4	9
Total	25	153	93	35	60	33	8	407

Appendix G-2c: Site 448155

Actual Accrual	Site Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	15	54	21	2	7	1	0	100
1 < 4	14	91	42	25	18	3	0	193
4 < 7	1	36	45	16	18	2	0	118
7 < 10	0	12	26	9	12	5	1	65
10 < 20	0	5	31	24	56	12	0	128
20 < 50	0	0	4	6	32	31	0	73
> 50	0	0	0	0	2	13	5	20
Total	30	198	169	82	145	67	6	697

Appendix G-2d: Site 494048

Actual Accrual	Site Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	14	72	18	8	5	1	0	118
1 < 4	13	107	37	9	17	3	1	187
4 < 7	3	36	26	4	9	1	0	79
7 < 10	0	9	20	7	10	3	0	49
10 < 20	0	7	22	18	18	3	2	70
20 < 50	0	0	3	5	19	14	0	41
> 50	0	0	0	0	2	3	3	8
Total	30	231	126	51	80	28	6	552

Appendix G-2e: Site 512786

Actual Accrual	Site Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	13	16	8	8	1	0	46
1 < 4	0	25	45	17	23	2	0	112
4 < 7	0	13	31	11	22	3	0	80
7 < 10	0	7	21	20	14	4	0	66
10 < 20	0	2	13	21	40	15	0	91
20 < 50	0	1	6	7	24	33	3	74
> 50	0	0	0	1	2	6	6	15
Total	0	61	132	85	133	64	9	484

Appendix G-2f: Site 560623

Actual Accrual	Site Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	64	14	3	0	0	0	0	81
1 < 4	26	32	8	2	0	1	0	69
4 < 7	1	7	4	1	0	0	0	13
7 < 10	0	5	3	0	0	0	0	8
10 < 20	0	1	2	0	3	1	0	7
20 < 50	0	0	1	0	0	0	0	1
> 50	0	0	0	0	0	0	0	0
Total	91	59	21	3	3	2	0	179

Appendix G-2g: Site 575415

Actual Accrual	Site Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	17	38	16	9	3	0	0	83
1 < 4	5	64	26	9	7	1	0	112
4 < 7	2	29	18	16	11	3	0	79
7 < 10	0	10	10	7	9	2	0	38
10 < 20	0	3	14	10	15	5	0	47
20 < 50	0	0	0	3	20	17	3	43
> 50	0	0	0	1	0	4	3	8
Total	24	144	84	55	65	32	6	410

Appendix G-2h: Site 598430

Actual Accrual	Site Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	22	60	11	4	0	1	0	98
1 < 4	15	79	17	4	1	0	0	116
4 < 7	0	32	13	4	2	0	0	51
7 < 10	0	10	7	7	3	0	0	27
10 < 20	0	4	8	6	5	0	0	23
20 < 50	0	0	0	1	6	3	0	10
> 50	0	0	0	0	0	1	0	1
Total	37	185	56	26	17	5	0	326

Appendix G-2i: Site 602591

Actual Accrual	Site Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	18	59	13	4	0	0	0	94
1 < 4	9	84	26	4	3	0	0	126
4 < 7	1	35	31	3	3	0	0	73
7 < 10	0	12	11	7	3	2	0	35
10 < 20	0	3	12	11	12	4	0	42
20 < 50	0	0	2	4	7	10	1	24
> 50	0	0	0	0	0	1	0	1
Total	28	193	95	33	28	17	1	395

Appendix G-2j: Site 689326

Actual Accrual	Site Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	0	0	0	0	0	0	0
1 < 4	5	38	51	17	17	4	0	132
4 < 7	0	13	25	24	21	4	1	88
7 < 10	0	0	15	14	11	4	0	44
10 < 20	0	2	19	20	44	19	1	105
20 < 50	0	0	1	8	18	31	5	63
> 50	0	0	0	0	2	12	10	24
Total	5	25	111	83	113	74	17	456

Appendix G-2k: Site 696337

Actual Accrual	Site Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	9	26	1	1	1	0	0	38
1 < 4	6	45	9	4	3	0	0	67
4 < 7	0	14	9	1	4	1	0	29
7 < 10	0	6	8	8	5	0	0	27
10 < 20	0	1	9	3	3	1	0	17
20 < 50	0	0	1	0	3	0	0	4
> 50	0	0	0	0	0	0	0	0
Total	15	92	37	17	19	2	0	182

Appendix G-2l: Site 714415

Actual Accrual	Site Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	13	47	13	5	2	0	0	80
1 < 4	2	77	45	16	8	4	0	152
4 < 7	0	35	32	14	11	0	0	92
7 < 10	0	10	12	23	13	3	1	62
10 < 20	0	9	9	13	31	10	0	72
20 < 50	0	1	3	2	21	28	0	55
> 50	0	0	0	0	2	5	9	16
Total	15	179	114	73	88	50	10	529

Appendix G-2m: Site 715532

Actual Accrual	Site Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	33	25	3	3	1	1	0	66
1 < 4	11	48	10	4	3	1	0	77
4 < 7	1	15	7	2	1	0	0	26
7 < 10	0	7	9	2	0	1	0	19
10 < 20	0	1	5	2	7	3	0	18
20 < 50	0	0	0	1	8	6	2	17
> 50	0	0	0	1	0	0	1	2
Total	45	96	34	15	20	12	3	225

Appendix G-2n: Site 846594

Actual Accrual	Site Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	16	38	5	2	1	2	0	64
1 < 4	3	76	27	2	3	4	0	115
4 < 7	0	20	11	5	1	2	0	39
7 < 10	0	5	3	2	1	0	2	13
10 < 20	0	2	8	10	13	5	0	38
20 < 50	0	0	2	3	9	7	2	23
> 50	0	0	0	1	0	0	1	2
Total	19	141	56	25	28	20	5	294

Appendix G-2o: Site 997056

Actual Accrual	Site Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	1	7	3	0	0	1	0	12
1 < 4	2	14	11	0	4	1	0	31
4 < 7	1	6	10	3	0	0	0	20
7 < 10	0	0	3	1	1	1	0	6
10 < 20	0	0	6	3	5	1	0	15
20 < 50	0	0	0	2	1	5	0	8
> 50	0	0	0	0	0	1	0	1
Total	4	27	33	9	11	9	0	93

Appendix G-2p: Site 998666

Actual Accrual	Site Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	16	31	7	1	0	0	0	55
1 < 4	3	37	14	3	1	0	0	58
4 < 7	0	19	8	1	1	0	0	29
7 < 10	0	5	6	0	3	0	0	14
10 < 20	0	3	6	4	7	4	0	24
20 < 50	0	0	0	2	4	4	2	12
> 50	0	0	0	0	0	1	0	1
Total	19	95	41	11	16	9	2	193

Appendix G-3: Disease Team Predicted versus Actual Accrual

Appendix G-3a: Site 104647

Actual Accrual	Disease Team Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	1	8	15	4	9	5	5	47
1 < 4	0	21	31	10	34	9	0	105
4 < 7	0	4	20	7	31	7	1	70
7 < 10	0	2	9	3	18	4	1	37
10 < 20	0	3	8	4	37	8	1	61
20 < 50	0	0	5	0	8	19	4	36
> 50	0	0	0	0	1	0	8	9
Total	1	38	88	28	138	52	20	365

Appendix G-3b: Site 173472

Actual Accrual	Disease Team Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	23	25	6	11	3	3	71
1 < 4	2	36	47	7	32	11	0	135
4 < 7	0	3	34	1	18	3	1	60
7 < 10	0	2	11	4	18	8	0	43
10 < 20	0	0	13	0	23	16	1	53
20 < 50	0	1	2	1	10	20	2	36
> 50	0	0	0	0	0	2	7	9
Total	2	65	132	19	112	63	14	407

Appendix G-3c: Site 448155

Actual Accrual	Disease Team Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	23	48	2	20	5	2	100
1 < 4	0	50	68	11	49	9	6	193
4 < 7	0	9	41	7	48	12	1	118
7 < 10	0	1	17	5	33	7	2	65
10 < 20	0	1	18	8	68	32	1	128
20 < 50	0	0	3	3	21	42	4	73
> 50	0	0	0	0	1	3	16	20
Total	0	84	195	36	240	110	32	697

Appendix G-3d: Site 494048

Actual Accrual	Disease Team Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	1	22	42	8	36	7	2	118
1 < 4	4	24	64	13	53	26	3	187
4 < 7	0	3	15	5	40	12	4	79
7 < 10	0	0	1	3	30	13	2	49
10 < 20	0	0	0	0	24	38	8	70
20 < 50	0	0	0	0	2	32	7	41
> 50	0	0	0	0	1	0	7	8
Total	5	49	122	29	186	128	33	552

Appendix G-3e: Site 512786

Actual Accrual	Disease Team Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	2	7	5	23	8	1	46
1 < 4	0	5	19	10	47	26	5	112
4 < 7	0	0	8	7	39	24	2	80
7 < 10	0	0	3	2	37	23	1	66
10 < 20	0	0	1	1	29	53	7	91
20 < 50	0	0	0	0	4	52	18	74
> 50	0	0	0	0	0	0	15	15
Total	0	7	38	25	179	186	49	484

Appendix G-3f: Site 560623

Actual Accrual	Disease Team Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	15	12	7	11	1	35	81
1 < 4	0	8	28	1	13	5	14	69
4 < 7	0	0	9	0	1	2	1	13
7 < 10	0	1	2	1	3	1	0	8
10 < 20	0	0	1	0	4	0	2	7
20 < 50	0	0	0	0	0	0	1	1
> 50	0	0	0	0	0	0	0	0
Total	0	24	52	9	32	9	53	179

Appendix G-3g: Site 575415

Actual Accrual	Disease Team Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	21	34	4	18	5	1	83
1 < 4	0	12	40	13	40	6	1	112
4 < 7	0	1	24	8	34	12	0	79
7 < 10	0	0	1	4	29	3	1	38
10 < 20	0	0	0	0	27	18	2	47
20 < 50	0	0	0	0	1	38	4	43
> 50	0	0	0	0	0	0	8	8
Total	0	34	99	29	149	82	17	410

Appendix G-3h: Site 598430

Actual Accrual	Disease Team Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	3	26	4	44	18	3	98
1 < 4	0	4	28	10	57	16	1	116
4 < 7	0	0	4	4	26	13	4	51
7 < 10	0	0	0	1	16	7	3	27
10 < 20	0	0	0	0	5	13	5	23
20 < 50	0	0	0	0	0	2	8	10
> 50	0	0	0	0	0	0	1	1
Total	0	7	58	19	148	69	25	326

Appendix G-3i: Site 602591

Actual Accrual	Disease Team Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	3	34	7	34	14	2	94
1 < 4	0	3	40	8	52	23	0	126
4 < 7	0	0	16	5	38	14	0	73
7 < 10	0	0	2	4	23	4	2	35
10 < 20	0	0	0	0	24	17	1	42
20 < 50	0	0	0	0	1	19	4	24
> 50	0	0	0	0	0	0	1	1
Total	0	6	92	24	172	91	10	395

Appendix G-3j: Site 689326

Actual Accrual	Disease Team Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	0	0	0	0	0	0	0
1 < 4	0	2	30	11	57	27	5	132
4 < 7	0	0	9	5	34	32	8	88
7 < 10	0	0	0	2	20	19	3	44
10 < 20	0	0	0	0	22	65	18	105
20 < 50	0	0	0	0	0	31	32	63
> 50	0	0	0	0	0	0	24	24
Total	0	2	39	18	133	174	90	456

Appendix G-3k: Site 696337

Actual Accrual	Disease Team Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	6	15	5	10	2	0	38
1 < 4	0	2	28	4	29	4	0	67
4 < 7	0	0	2	2	21	4	0	29
7 < 10	0	0	2	2	19	4	0	27
10 < 20	0	0	3	1	7	6	0	17
20 < 50	0	0	0	0	2	2	0	4
> 50	0	0	0	0	0	0	0	0
Total	0	8	50	14	88	22	0	182

Appendix G-3l: Site 714415

Actual Accrual	Disease Team Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	22	25	6	22	5	0	80
1 < 4	0	22	62	12	41	14	1	152
4 < 7	0	5	30	14	31	10	2	92
7 < 10	0	0	6	7	35	13	1	62
10 < 20	0	0	3	4	40	24	1	72
20 < 50	0	0	0	1	5	44	5	55
> 50	1	0	0	0	0	3	12	16
Total	1	49	126	44	174	113	22	529

Appendix G-3m: Site 715532

Actual Accrual	Disease Team Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	17	26	5	13	3	2	66
1 < 4	0	9	21	13	24	7	3	77
4 < 7	0	1	3	3	13	5	1	26
7 < 10	0	0	0	1	11	5	2	19
10 < 20	0	0	0	0	1	9	8	18
20 < 50	0	0	0	0	3	7	7	17
> 50	0	0	0	0	0	0	2	2
Total	0	27	50	22	65	36	25	225

Appendix G-3n: Site 846594

Actual Accrual	Disease Team Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	13	25	5	8	11	2	64
1 < 4	0	5	32	9	40	12	17	115
4 < 7	0	0	3	4	17	12	3	39
7 < 10	0	0	0	0	5	4	4	13
10 < 20	0	0	0	1	4	22	11	38
20 < 50	0	0	0	0	0	10	13	23
> 50	0	0	0	0	0	0	2	2
Total	0	18	60	19	74	71	52	294

Appendix G-3o: Site 997056

Actual Accrual	Disease Team Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	0	3	1	7	1	0	12
1 < 4	0	2	8	7	11	3	0	31
4 < 7	0	0	4	4	12	0	0	20
7 < 10	0	0	0	0	5	1	0	6
10 < 20	0	0	0	0	6	9	0	15
20 < 50	0	0	0	0	0	6	2	8
> 50	0	0	0	0	0	0	1	1
Total	0	2	15	12	41	20	3	93

Appendix G-3p: Site 998666

Actual Accrual	Disease Team Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	12	19	2	16	6	0	55
1 < 4	0	10	21	7	14	4	2	58
4 < 7	0	1	11	1	14	2	0	29
7 < 10	0	2	1	0	9	2	0	14
10 < 20	0	1	4	1	10	5	3	24
20 < 50	0	0	1	0	4	7	0	12
> 50	0	0	0	0	1	0	0	1
Total	0	26	57	11	68	26	5	193

Appendix G-4: Overall Prediction Model versus Actual Accrual

Appendix G-4a: Site 104647

Actual Accrual	Overall Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	2	19	13	2	5	3	3	47
1 < 4	1	48	29	20	7	0	0	105
4 < 7	0	21	31	11	7	0	0	70
7 < 10	0	14	10	7	5	1	0	37
10 < 20	0	5	23	10	20	3	0	61
20 < 50	0	2	6	2	19	7	0	36
> 50	0	0	0	2	0	4	3	9
Total	3	109	112	54	63	18	6	365

Appendix G-4b: Site 173472

Actual Accrual	Overall Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	4	41	13	9	2	0	2	71
1 < 4	3	67	42	7	13	3	0	135
4 < 7	0	18	22	11	6	3	0	60
7 < 10	0	6	17	11	6	3	0	43
10 < 20	0	2	12	15	18	6	0	53
20 < 50	0	0	4	2	15	14	1	36
> 50	0	0	0	1	0	5	3	9
Total	7	134	110	56	60	34	6	407

Appendix G-4c: Site 448155

Actual Accrual	Overall Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	7	53	30	6	4	0	0	100
1 < 4	5	91	66	16	14	0	1	193
4 < 7	0	45	45	14	14	0	0	118
7 < 10	0	14	25	15	8	3	0	65
10 < 20	0	14	44	26	39	5	0	128
20 < 50	0	2	13	11	28	18	1	73
> 50	0	0	0	0	6	13	1	20
Total	12	219	223	88	113	39	3	697

Appendix G-4d: Site 494048

Actual Accrual	Overall Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	2	50	37	13	12	4	0	118
1 < 4	2	79	69	19	12	5	1	187
4 < 7	0	22	30	12	13	2	0	79
7 < 10	0	3	17	12	14	3	0	49
10 < 20	0	6	15	18	28	3	0	70
20 < 50	0	0	3	7	20	11	0	41
> 50	0	0	0	0	1	6	1	8
Total	4	160	171	81	100	34	2	552

Appendix G-4e: Site 512786

Actual Accrual	Overall Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	3	11	16	7	8	1	0	46
1 < 4	1	41	37	11	19	3	0	112
4 < 7	1	23	31	9	13	3	0	80
7 < 10	0	12	27	9	12	6	0	66
10 < 20	0	10	21	23	22	14	1	91
20 < 50	0	3	9	9	15	36	2	74
> 50	0	0	1	1	3	3	7	15
Total	5	100	142	69	92	66	10	484

Appendix G-4f: Site 560623

Actual Accrual	Overall Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	2	50	22	3	4	0	0	81
1 < 4	0	21	26	10	8	4	0	69
4 < 7	0	3	2	7	0	0	1	13
7 < 10	0	1	2	4	1	0	0	8
10 < 20	0	0	0	5	0	2	0	7
20 < 50	0	0	0	0	0	1	0	1
> 50	0	0	0	0	0	0	0	0
Total	2	75	52	29	13	7	1	179

Appendix G-4g: Site 575415

Actual Accrual	Overall Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	10	35	20	10	7	1	0	83
1 < 4	0	50	35	14	12	1	0	112
4 < 7	0	24	30	11	11	3	0	79
7 < 10	0	8	14	5	9	2	0	38
10 < 20	0	4	13	11	16	3	0	47
20 < 50	0	0	1	6	14	20	2	43
> 50	0	0	0	0	1	4	3	8
Total	10	121	113	57	70	34	5	410

Appendix G-4h: Site 598430

Actual Accrual	Overall Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	6	46	28	8	8	1	1	98
1 < 4	3	52	43	12	3	3	0	116
4 < 7	0	16	19	8	7	1	0	51
7 < 10	0	4	7	5	8	3	0	27
10 < 20	0	0	10	2	8	3	0	23
20 < 50	0	0	0	0	4	4	2	10
> 50	0	0	0	0	0	1	0	1
Total	9	118	107	35	38	16	3	326

Appendix G-4i: Site 602591

Actual Accrual	Overall Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	6	48	26	10	4	0	0	94
1 < 4	3	65	38	12	8	0	0	126
4 < 7	0	24	27	14	7	1	0	73
7 < 10	0	3	15	10	6	1	0	35
10 < 20	0	3	12	5	17	5	0	42
20 < 50	0	0	1	3	6	14	0	24
> 50	0	0	0	0	1	0	0	1
Total	9	143	119	54	49	21	0	395

Appendix G-4j: Site 689326

Actual Accrual	Overall Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	0	0	0	0	0	0	0
1 < 4	3	68	43	13	4	1	0	132
4 < 7	1	23	42	15	3	3	1	88
7 < 10	0	11	12	10	10	1	0	44
10 < 20	0	17	32	21	27	8	0	105
20 < 50	0	2	14	6	23	18	0	63
> 50	0	0	1	1	9	9	4	24
Total	4	121	144	66	76	40	5	456

Appendix G-4k: Site 696337

Actual Accrual	Overall Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	21	14	1	2	0	0	38
1 < 4	1	27	25	9	5	0	0	67
4 < 7	0	11	10	7	0	1	0	29
7 < 10	0	6	17	4	0	0	0	27
10 < 20	0	4	6	3	4	0	0	17
20 < 50	0	1	2	1	0	0	0	4
> 50	0	0	0	0	0	0	0	0
Total	1	70	74	25	11	1	0	182

Appendix G-4l: Site 714415

Actual Accrual	Overall Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	8	50	16	4	2	0	0	80
1 < 4	4	64	60	11	11	2	0	152
4 < 7	0	34	35	14	9	0	0	92
7 < 10	0	8	19	21	11	3	0	62
10 < 20	0	6	18	16	26	6	0	72
20 < 50	0	1	5	9	18	22	0	55
> 50	0	0	0	0	1	6	9	16
Total	12	163	153	75	78	39	9	529

Appendix G-4m: Site 715532

Actual Accrual	Overall Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	3	37	17	6	3	0	0	66
1 < 4	0	22	31	15	6	3	0	77
4 < 7	0	8	9	6	3	0	0	26
7 < 10	0	3	5	8	3	0	0	19
10 < 20	0	0	6	3	5	3	1	18
20 < 50	0	0	0	1	5	11	0	17
> 50	0	0	0	1	0	0	1	2
Total	3	70	68	0	25	17	2	225

Appendix G-4n: Site 846594

Actual Accrual	Overall Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	10	39	7	3	3	1	1	64
1 < 4	5	42	36	23	7	2	0	115
4 < 7	0	11	15	9	3	0	1	39
7 < 10	0	2	4	2	4	1	0	13
10 < 20	0	2	5	9	13	9	0	38
20 < 50	0	1	0	2	9	11	0	23
> 50	0	1	0	0	0	0	1	2
Total	15	98	67	48	39	24	3	294

Appendix G-4o: Site 997056

Actual Accrual	Overall Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	1	5	4	1	1	0	0	12
1 < 4	2	13	11	3	2	0	0	31
4 < 7	0	7	10	2	1	0	0	20
7 < 10	0	0	4	1	1	0	0	6
10 < 20	0	1	2	7	4	1	0	15
20 < 50	0	0	2	0	1	5	0	8
> 50	0	0	0	0	0	1	0	1
Total	3	26	33	14	10	7	0	93

Appendix G-4p: Site 998666

Actual Accrual	Overall Model Predicted Accrual							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	1	27	17	9	1	0	0	55
1 < 4	1	28	22	3	4	0	0	58
4 < 7	1	9	13	5	1	0	0	29
7 < 10	0	6	6	2	0	0	0	14
10 < 20	1	1	4	7	10	1	0	24
20 < 50	0	1	0	2	4	5	0	12
> 50	0	0	0	0	0	1	0	1
Total	4	72	62	28	20	7	0	193

Appendix G-5: Overall Prediction Model: Random Effects versus Actual Accrual

Appendix G-5a: Site 104647

Actual Accrual	Overall Model Predicted Accrual: Random Effects							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	2	19	12	3	5	3	3	47
1 < 4	0	44	31	21	9	0	0	105
4 < 7	0	14	33	16	7	0	0	70
7 < 10	0	1	2	12	7	1	0	37
10 < 20	0	4	23	9	22	3	0	61
20 < 50	0	0	7	2	17	10	0	36
> 50	0	0	0	2	0	4	3	9
Total	2	93	118	60	65	21	6	365

Appendix G-5b: Site 173472

Actual Accrual	Overall Model Predicted Accrual: Random Effects							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	4	41	13	8	3	0	2	71
1 < 4	3	62	40	13	14	3	0	135
4 < 7	0	15	23	11	8	3	0	60
7 < 10	0	5	17	11	7	3	0	43
10 < 20	0	1	12	13	21	6	0	53
20 < 50	0	0	3	3	13	16	1	36
> 50	0	0	0	0	1	5	3	9
Total	7	124	108	59	67	36	6	407

Appendix G-5c: Site 448155

Actual Accrual	Overall Model Predicted Accrual: Random Effects							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	7	52	28	7	6	0	0	100
1 < 4	5	84	72	17	13	1	1	193
4 < 7	0	43	44	17	13	1	0	118
7 < 10	0	13	22	19	8	3	0	65
10 < 20	0	13	40	28	42	5	0	128
20 < 50	0	1	12	13	27	19	1	73
> 50	0	0	0	0	5	14	1	20
Total	12	206	218	101	114	43	3	697

Appendix G-5d: Site 494048

Actual Accrual	Overall Model Predicted Accrual: Random Effects							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	2	48	36	14	14	4	0	118
1 < 4	2	75	70	22	11	6	1	187
4 < 7	0	15	37	10	14	3	0	79
7 < 10	0	1	17	13	14	4	0	49
10 < 20	0	5	14	17	31	3	0	70
20 < 50	0	0	3	6	20	12	0	41
> 50	0	0	0	0	1	6	1	8
Total	4	144	177	82	105	38	2	552

Appendix G-5e: Site 512786

Actual Accrual	Overall Model Predicted Accrual: Random Effects							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	3	11	14	8	9	1	0	46
1 < 4	1	37	39	13	19	3	0	112
4 < 7	1	21	32	9	13	4	0	80
7 < 10	0	11	26	11	11	7	0	66
10 < 20	0	9	21	22	24	14	1	91
20 < 50	0	2	10	6	18	36	2	74
> 50	0	0	1	1	2	4	7	15
Total	5	91	143	70	96	69	10	484

Appendix G-5f: Site 560623

Actual Accrual	Overall Model Predicted Accrual: Random Effects							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	41	36	4	0	0	0	0	81
1 < 4	11	51	3	1	3	0	0	69
4 < 7	1	11	0	0	1	0	0	13
7 < 10	0	7	1	0	0	0	0	8
10 < 20	0	5	0	0	2	0	0	7
20 < 50	0	0	0	1	0	0	0	1
> 50	0	0	0	0	0	0	0	0
Total	53	110	8	2	6	0	0	179

Appendix G-5g: Site 575415

Actual Accrual	Overall Model Predicted Accrual: Random Effects							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	8	35	21	9	9	1	0	83
1 < 4	0	47	35	15	14	1	0	112
4 < 7	0	20	31	12	13	3	0	79
7 < 10	0	7	15	5	9	2	0	38
10 < 20	0	4	12	9	19	3	0	47
20 < 50	0	0	1	6	12	22	2	43
> 50	0	0	0	0	1	4	3	8
Total	8	113	115	56	77	36	5	410

Appendix G-5h: Site 598430

Actual Accrual	Overall Model Predicted Accrual: Random Effects							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	6	42	32	5	11	1	1	98
1 < 4	2	45	43	20	3	3	0	116
4 < 7	0	15	19	7	9	1	0	51
7 < 10	0	3	7	5	9	3	0	27
10 < 20	0	0	9	2	8	4	0	23
20 < 50	0	0	0	0	4	3	3	10
> 50	0	0	0	0	0	1	0	1
Total	8	105	110	39	44	16	4	326

Appendix G-5i: Site 602591

Actual Accrual	Overall Model Predicted Accrual: Random Effects							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	6	46	27	10	5	0	0	94
1 < 4	2	56	46	12	10	0	0	126
4 < 7	0	22	27	15	8	1	0	73
7 < 10	0	2	16	10	6	1	0	35
10 < 20	0	2	11	6	18	5	0	42
20 < 50	0	0	1	2	7	14	0	24
> 50	0	0	0	0	1	0	0	1
Total	8	128	128	55	55	21	0	395

Appendix G-5j: Site 689326

Actual Accrual	Overall Model Predicted Accrual: Random Effects							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	0	0	0	0	0	0	0
1 < 4	3	61	48	13	6	1	0	132
4 < 7	0	20	44	15	5	3	1	88
7 < 10	0	9	12	10	12	1	0	44
10 < 20	0	12	31	26	27	9	0	105
20 < 50	0	2	14	5	24	18	0	63
> 50	0	0	1	0	10	8	5	24
Total	3	104	150	69	84	40	6	456

Appendix G-5k: Site 696337

Actual Accrual	Overall Model Predicted Accrual: Random Effects							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	0	24	12	0	2	0	0	38
1 < 4	1	26	27	8	5	0	0	67
4 < 7	0	13	8	7	0	1	0	29
7 < 10	0	7	17	3	0	0	0	27
10 < 20	0	4	6	4	3	0	0	17
20 < 50	0	1	2	1	0	0	0	4
> 50	0	0	0	0	0	0	0	0
Total	1	75	72	23	10	1	0	182

Appendix G-5l: Site 714415

Actual Accrual	Overall Model Predicted Accrual: Random Effects							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	8	49	16	5	2	0	0	80
1 < 4	4	56	60	19	11	2	0	153
4 < 7	0	31	37	11	13	0	0	92
7 < 10	0	7	18	23	11	3	0	62
10 < 20	0	5	18	16	26	7	0	72
20 < 50	0	1	4	8	17	25	0	55
> 50	0	0	0	0	1	6	9	16
Total	12	149	153	82	81	43	9	529

Appendix G-5m: Site 715532

Actual Accrual	Overall Model Predicted Accrual: Random Effects							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	9	43	11	2	1	0	0	66
1 < 4	4	47	18	4	4	0	0	77
4 < 7	0	13	11	1	1	0	0	26
7 < 10	0	8	8	1	2	0	0	19
10 < 20	0	4	6	1	5	2	0	18
20 < 50	0	0	1	1	10	5	0	17
> 50	0	0	1	0	0	1	0	2
Total	16	115	56	10	23	8	0	225

Appendix G-5n: Site 846594

Actual Accrual	Overall Model Predicted Accrual: Random Effects							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	10	36	10	2	4	1	1	64
1 < 4	5	39	37	23	9	2	0	115
4 < 7	0	11	14	9	4	0	1	39
7 < 10	0	2	4	1	5	1	0	13
10 < 20	0	2	5	8	13	10	0	38
20 < 50	0	1	0	2	9	11	0	23
> 50	0	1	0	0	0	0	1	2
Total	15	92	70	45	44	25	3	294

Appendix G-5o: Site 997056

Actual Accrual	Overall Model Predicted Accrual: Random Effects							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	1	4	5	1	1	0	0	12
1 < 4	1	14	11	2	3	0	0	31
4 < 7	0	6	11	2	1	0	0	20
7 < 10	0	0	4	1	1	0	0	6
10 < 20	0	1	2	5	6	1	0	15
20 < 50	0	0	2	0	1	5	0	8
> 50	0	0	0	0	0	1	0	1
Total	2	25	35	11	13	7	0	93

Appendix G-5p: Site 998666

Actual Accrual	Overall Model Predicted Accrual: Random Effects							Total
	< 1	1 < 4	4 < 7	7 < 10	10<20	20<50	> 50	
< 1	1	33	19	2	0	0	0	55
1	1	35	16	4	2	0	0	58
1	1	16	8	4	0	0	0	29
7 < 10	0	6	7	1	0	0	0	14
10 < 20	1	1	8	8	6	0	0	24
20 < 50	0	1	0	3	6	2	0	12
> 50	0	0	0	0	0	1	0	1
Total	4	92	58	22	14	3	0	193

REFERENCES

1. IOM (Institute of Medicine). 2010. Transforming Clinical Research in the United States: Challenges and Opportunities: Workshop Summary. Washington, DC: The National Academies Press.
2. IOM (Institute of Medicine). 2010. A National Cancer Clinical Trials System for the 21st Century: Reinvigorating the NCI Cooperative Group Program. Washington, DC: The National Academies Press.
3. Al-Shahi Salman R, Beller E, Kagan J, Hemminki E, Phillips RS, Savuescu J, Macleod M, Wisely J, Chalmers I. Increasing value and reducing waste in biomedical research regulation and management. *Lancet*. 2014; 383(9913):176-85.
4. Assistant Secretary for Legislation (ASL), Department of Health & Human Services. Testimony on the Value of Clinical Research by Harold Varmus, M.D [Internet]. Bethesda, MD: Senate Comm. on Labor and Human Resources, Subcommittee on Public Health and Safety. 1997 Oct 9 [cited 2015 Jun 23]. Available from: www.hhs.gov/asl/testify/t971009a.html.
5. DiMasi JA. Cost of Developing a New Drug [briefing]. Boston (MA): Tufts Center for the Study of Drug Development; 2014 Nov 18 [cited 2015 Jan 14]. Available from: http://csdd.tufts.edu/files/uploads/Tufts_CSDD_briefing_on_RD_cost_study_-_Nov_18,_2014..pdf.
6. Chalmers I, Bracken MB, Djulbegovic B, Garattini S, Grant J, Gulmezoglu AM, Howells DW, Ioannidis JPA, Oliver S. How to increase value and reduce waste when research priorities are set. *Lancet*. 2014; 383(9912):156-65.

7. Kolata G. Grant systems leads researchers to play it safe [Internet]. New York Times. 2009 Jun 27 [updated 2009 Jul 5; cited 2015 Jun 23]. Available from: http://www.nytimes.com/2009/06/28/health/research/28cancer.html?pagewanted=all&_r=1&_r=1&_r=1
8. Centers for Disease Control and Prevention (CDC). Leading Causes of Death [Internet]. 2015 Feb 6 [cited 2015 Jun 23]. Available from: <http://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm>.
9. American Society of Oncology. The State of Cancer Care in America, 2014: A Report by the American Society of Clinical Oncology. J Oncol Pract. 2014; 10(2):119-42.
10. Sertkaya A, Birkenbach A, Eyraud J (Eastern Research Group, Inc., Lexington, MA). Examination of Clinical Trial Costs and Barriers for Drug Development. Final report 25 Jul 2014. Washington, DC: Department of Health and Human Services, Assistant Secretary of Planning and Evaluation (ASPE). 92p. Task Order No. HHSP23337007T. Contract HHSP23320095634WC.
11. National Cancer Institute. What is Cancer? [Internet]. Updated 2015 Feb 9 [cited 2015 Jun 23]. Available from: <http://www.cancer.gov/about-cancer/what-is-cancer>.
12. The White House. FACT SHEET: President Obama's Precision Medicine Initiative [Internet]. Washington, DC: White House; 2015 Jan 30 [cited 2015 Jun 23]. Accessed from: <https://www.whitehouse.gov/the-press-office/2015/01/30/fact-sheet-president-obama-s-precision-medicine-initiative>.

13. National Cancer Institute. NCI and the Precision Medicine Initiative [Internet]. Updated 2015 Jun 5 [cited 2015 Jun 23]. Available from: <http://www.cancer.gov/research/key-initiatives/precision-medicine>.
14. The White House. Remarks of President Barak Obama - State of the Union Address as Delivered [Internet]. Washington, DC: White House; 2016 Jan 13 [cited 2016 Jan 30]. Accessed from: <https://www.whitehouse.gov/the-press-office/2016/01/12/remarks-president-barack-obama-%E2%80%93-prepared-delivery-state-union-address>.
15. DiMasi JA, Grabowski HG. Economics of New Oncology Drug Development. *J Clin Oncol*. 2007; 25(2):209-16.
16. Heller C, de Melo-Martin I. Clinical and Translational Science Awards: Can They Increase the Efficiency and Speed of Clinical and Translational Research? *Acad Med*. 2009; 84(4):424-32.
17. Campo RA, Seguire ES, inventors; Medidata Ft, Inc., assignee. Protocol complexity analyzer. United States patent application US WO 2010030530 A2. 2008 Sep 11.
18. Dilts DM, Cheng SK, Crites JS, Sandler AB, Doroshow JH. Phase III Clinical Trial Development: A Process of Chutes and Ladders. *Clin Cancer Res*. 2010, 16(22):5381-89.
19. Dilts DM, Sandler AB, Baker M, Cheng SK, George SL, Karas KS, McGuire S, Menon GS, Reusch J, Sawyer D, Scoggins M, Wu A, Zhou K, Schilsky RL. Processes to Activate Phase III Clinical Trials in a Cooperative Oncology Group: The Case of Cancer and Leukemia Group B. *J Clin Oncol*. 2006; 24(28):4553-7.

20. Dilts DM, Sandler AB. Invisible Barriers to Clinical Trials: The Impact of Structural, Infrastructural, and Procedural Barriers to Opening Oncology Clinical Trials. *J Clin Oncol*. 2006; 24(28):4545-52.
21. IOM (Institute of Medicine). 2012. *Envisioning a Transformed Clinical Trials Enterprise in the United States: Establishing an Agenda for 2020*. Washington, DC: The National Academies Press.
22. Denny TL, Yasko JM, Collyar DE, Katz ML, Krasnov CL, Borwhat MJ, Battershell A, George SL. Managing Accrual in Cooperative Group Clinical Trials. *J Clin Oncol* 2004; 22(15):2997-3002.
23. Schroen AT, Petroni GR, Wang H, Gray R, Wang XF, Cronin W, Sargent DJ, Benedetti J, Wickerham DL, Djulbegovic B, Slingluff CL Jr. Preliminary evaluation of factors associated with premature trial closure and feasibility of accrual benchmarks in phase III oncology trials. *Clin Trials*. 2010, 7(4):312-21.
24. Cheng SK, Dietrich MS, Dilts DM. A Sense of Urgency: Evaluating the Link between Clinical Trial Development Time and the Accrual Performance of Cancer Therapy Evaluation Program (NCI-CTEP) Sponsored Studies. *Clin Cancer Res*. 2010, 16(22):5557-63.
25. Dilts DM, Sandler AB, Cheng S, Crites L, Ferranti A, Bookman MA, Thomas JP, Ostroff J. Accrual to clinical trials at selected comprehensive cancer centers. *J Clin Oncol*. 2008; 26(suppl):abstract 6543.
26. Durivage HJ, Bridges KD. Clinical trial metrics: Protocol Performance and Resource Utilization from 14 Cancer Centers. *J Clin Oncol*. 2009; 27(15s):S6557.

27. Ioannidis JPA, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, Schulz KF, Tibshirani R. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet*. 2014; 383(9912):166-75.
28. Warlow C. Over-regulation of clinical research: a threat to public health. *Clin Med*. 2005; 5(1):33-8.
29. Kramer JM, Smith PB, Califf RM. Impediments to Clinical Research in the United States. *Clin Pharmacol Ther*. 2012; 91(3):535-41.
30. Harinstein ME, Butler J, Greene SJ, Fonarow GC, Stockbridge NL, O'Connor CM, Pfeffer MA, Mehra MR, Solomon SD, Yancy CW, Fiuzat M, Mentz RJ, Collins SP, McMurray, Vaduganathan M, Dunnmon PM, Rosano GMC, Dinh W, Misselwitz F, Bonow RO, Gheorghide M. Site selection for heart failure clinical trials in the USA. *Heart Fail Rev*. 2015; 20(4):375-83.
31. Buonansegna E, Salomo S, Maier AM, Li-Ying J. Pharmaceutical new product development: why do clinical trials fail? *R&D Manage*. 2014; 44(2):189-202.
32. Koski G, Tobin MF, Whalen M. The Synergy of the Whole: Building a Global System for Clinical Trials to Accelerate Medicines Development. *Clin Ther*. 2014; 36(10):1356-70.
33. Comis RL, Crowley J. Baseline Study of Patient Accrual Onto Publicly Sponsored U.S. Cancer Clinical Trials: An analysis conducted for the Global Access Project of the National Patient Advocate Foundation [Internet]. 2006 Feb 1 [cited 2015 Jun 18]. Available from: www.npaf.org/sites/all/themes/NPAF/images/pdfs/cccg.pdf.

34. IOM (Institute of Medicine). 2012. The Clinical Trials Enterprise in the United States: A Call for Disruptive Innovation. Washington, DC: The National Academies Press.
35. Dobson A, Ko K, Gallese P, Lovett III EJ, Collins C, Butterfield N. A guidance document for implementing effective cancer clinical trials executive summary, Version 1.2 [Internet]. Washington, DC: C-Change; 2007 [cited 30 Jun 2015]. 52p. Available from: <http://c-change.together.org/Websites/cchange/images/Publications%20and%20Reports/ClinicalTrialsSuccess.pdf>.
36. Jenkins V, Fallowfield L. Reasons for accepting or declining to participate in randomized clinical trials for cancer therapy. *Br J Cancer*. 2000; 82(11):1783-8.
37. The Lipid Research Clinics Program. Participant recruitment to the coronary primary prevention trial. *J Chronic Dis*. 1983; 36(6):461-5.
38. The Collaborative Ocular Melanoma Study Group. Sociodemographic and Clinical Predictors of Participation in Two Randomized Trials: Findings from the Collaborative Ocular Melanoma Study COMS Report No. 7. *Control Clin Trials*. 2001; 22(5):526-37.
39. National Institutes of Health. Clinical and Translational Science Awards (CTSA) [Internet]. Bethesda (MD). 2015 Jun 8 [cited 2015 Jun 30]. Available from: <http://www.ncats.nih.gov/ctsa>.
40. IOM (Institute of Medicine). 2012. Transforming the Economics of Clinical Trials. Institute of Medicine. Washington, DC: The National Academies Press.
41. Griffith TL, Sawyer JE. Research Team Design and Management for Centralized R&D. *IEEE T Eng Manage*. 2010; 57(2):211-24.

42. Bertolotti F, Mattarelli, Vignoli M, Macri. Exploring the relationship between multiple team membership and team performance: The role of social networks and collaborative technology. *Res Policy*. 2015; 44(4):911-24.
43. Schultz C, Schreyoegg J, von Reitzenstein C. The moderating role of internal and external resources on the performance effect of multitasking: Evidence from the R&D performance of surgeons. *Res Policy*. 2013; 42(8):1356-65.
44. Blindenbach-Driessen F. The (In)Effectiveness of Cross-Functional Innovation Teams: The Moderating Role of Organizational Context. *IEEE T Eng Manage*. 2015; 62(1):29-38.
45. Cooper RG, Edgett SJ, Klienschmidt EJ. Optimizing the Stage-Gate® Process: What Best Practice Companies are Doing (Part Two) [Internet]. *Research Technology Management; Industrial Research Institute, Inc.* 2002; 25(5). 11p. Available from: http://www.stage-gate.com/resources_stage-gate_wp15.php
46. Department of Health and Human Services. Enterprise Performance Life Cycle [Internet]. [cited 2015 Jun 29]. Accessed from: <http://www.hhs.gov/ocio/eplc/index.html>.
47. Barnard KD, Dent L, Cook A. A systematic review of models to predict recruitment to multicentre clinical trials. *BMC Med Res Methodol*. 2010; 10:63-9.
48. Korn EL, Freidlin B, Mooney M, Abrams JS. Accrual Experience of National Cancer Institute Cooperative Group Phase III Trials Activated From 2000 to 2007. *J Clin Oncol*. 2010; 28(35):5197-201.

49. Macleod MR, Michie S, Roberts I, Dirnagl U, Chalmers I, Ioannidis JPA, Al-Shai Salman R, Chan AW, Glasziou P. Biomedical research: increasing value, reducing waste. *Lancet*. 2014; 383(9912):101-4.
50. Chan AW, Song F, Vickers A, Jefferson T, Dickersin K, Gotzsche PC, Krumholz HM, Gherzi D, van der Worp HB. Increasing value and reducing waste: addressing inaccessible research. *Lancet*. 2014; 383(9913):257-66.
51. Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, Michie S, Michie S, Moher D, Wager. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet*. 2014; 383(9913):267-76.
52. Morris ZS, Wooding S, Grant J. The answer is 17 years, what is the question: understand time lags in translational research. *J R Soc Med*. 2011; 104(12): 510-20.
53. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet*. 2009; 374(9683):86-9.
54. Prinz F, Schlange T, Asadullah K. Believe it or not: How much can we rely on published data on potential drug targets? *Nat Rev Drug Discov*. 2011; 10(9):712-3.
55. Begley CG, Ellis LM. Raise standards for preclinical cancer research: C. Glenn Begley and Lee M. Ellis propose how methods, publications and incentives must change if patients are to benefit. *Nature* 2012; 483(7391):531-3.
56. Tunis SR, Stryer DB, Clancy CM. Practical Clinical Trials: Increasing the Value of Clinical Research for Decision Making in Clinical and Health Policy. *JAMA*. 2003; 290(12):1624-32.

57. IOM (Institute of Medicine). 2013. The CTSA Program at NIH: Opportunities for Advancing Clinical and Translational Research. Washington, DC: The National Academies of Science.
58. American Society for Clinical Oncology. Federally Funded Cancer Research [Internet]. Alexandria, VA; 2015. Available from: <http://www.asco.org/advocacy/federally-funded-cancer-research>.
59. Piantadosi S, Patterson B. A Method for Predicting Accrual, Cost, and Paper Flow in Clinical Trials. *Control Clin Trials*. 1987; 8(3):205-15.
60. Emanuel E, Schnipper L, Kamin D, Levinson J, Lichter A. The costs of conducting clinical research. *J Clin Oncol*. 2003; 21(22):4145-50.
61. Steensma DP, Kantarjian HM. Impact of Cancer Research Bureaucracy on Innovation, Costs, and Patient Care. *J Clin Oncol*. 2014; 32(5):376-8.
62. Berndt ER, Cockburn IM (School of Management, Boston University, Cambridge, MA). Price Indexes for Clinical Trial Research: A Feasibility Study. Draft report 26 Dec 2012. National Bureau of Economic Research: 2012. 27p. Award Number 0915674.
63. Abrams JS, Mooney M, Zweibel J, Friedman S. Improving the Protocol Implementation Process: The National Cancer Institute's Response. *J Clin Oncol*. 2014; 32(23):2515-6.
64. Abrams JS, Mooney MM, Zwiebel JA, Korn EL, Friedman SH, Finnigan SR, Schettino PR, Denicoff AM, Kruhm MG, Montello M, Misra RR, Ansher SS, DiPiazza KJ, Souhan EM, Wickerham DL, Giantonio BJ, O'Donnell RT, Sullivan DM, Soto NI, Fleming GF, Prindiville SA, Petryshyn RA, Hautala JA, Grad O, Zuckerman BL, Meyer RM, Yao JC, Baker LA, Buckner JC, Hortobagyi GN, Doroshow JH. Implementation of timeline

- reforms speeds initiation of National Cancer Institute-sponsored trials. *J Natl Cancer Inst.* 2013; 105(13):954-9.
65. Mandrekar SJ, Dahlbeg SE, Simon R. Improving Clinical Trial Efficiency: Thinking outside the Box [Internet]. American Society of Clinical Oncology Educational Book. 2015; 35:e141-7. Available from: <http://meetinglibrary.asco.org/content/11500141-156>.
 66. ECOG-ACRIN Cancer Research Group. NCI-MATCH/EAY 131 [Internet]. [cited 2015 Oct 3]. Available at <http://ecog-acrin.org/trials/nci-match-eay131>.
 67. Quigley NR, Tesluk PE, Locke EA, Bartol KM. A multilevel investigation of the motivational mechanisms underlying knowledge sharing and performance. *Organ Sci.* 2007; 18(1):71-88.
 68. National Cancer Institute. PAR-13-386. Cancer Center Support Grants (CCSGs) for NCI-designated Cancer Centers (P30). 2013 Nov 7 [cited 2015 Aug 24]. Available from: <http://grants.nih.gov/grants/guide/pa-files/PAR-13-386.html>.
 69. Ning N, Yan J, Xie X, Gerber DE. Impact of NCI-Mandated Scientific Review on Protocol Development and Content. *J Natl Compr Canc Netw.* 2015; 13(4):409-16.
 70. Schoenberger JA. Recruitment in the Coronary Drug Project and the Aspirin Myocardial Infarction Study. *Clin Pharmacol Ther.* 1979; 25(5, part 2):681-4.
 71. Schroen AT, Petroni GR, Wang H, Thielen MJ, Sargent D, Benedetti JK, Cronin WM, Wickerham DL, Wang XF, Gray R, Cohn WF, Slingluff CL, Djulbegovic B. Challenges to accrual predictions to phase III cancer clinical trials: a survey of study chairs and lead statisticians of 248 NCI-sponsored trials. *Clin Trials.* 2011; 8(5):591-600.

72. American Society for Clinical Oncology. One in Five Clinical Trials for Adults with Cancer Never Finish – New Study Examines the Reasons [Internet]. Genitourinary Cancers Symposium. 2014 Jan 28 [cited 2015 Jun 30]. Accessed from <http://www.cancer.net/one-five-clinical-trials-adults-cancer-never-finish-%E2%80%93-new-study-examines-reasons>.
73. Rosas SR, Schouten JT, Dixon D, Varghese S, Cope MT, Marci J, Kagan JM. Evaluating protocol lifecycle time intervals in HIV/AIDS clinical trials. *Clin Trials*. 2014; 11(5):553-9.
74. The National Commission for the Protection of Human Subjects. Ethical Principles and Guidelines for the Protection of Human Subjects of Research [Internet]. 1979 Apr 18 [cited 2015 May 13]. Available from: <http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html>.
75. Buchwald H, Matts JP, Hansen BJ, Long JM, Fitch LL. Program on Surgical Control of the Hyperlipidemias (POSCH): Recruitment Experience. *Control Clin Trials*. 1987; 8(4 Suppl):94S-104S.
76. Yeomans-Kinney A, Vernon SW, Frankowski RF, Weber DM, Bitsura JM, Vogel VG. Factors related to enrollment in the Breast Cancer Prevention Trial at a comprehensive cancer center during the first year of recruitment. *Cancer*. 1995; 76(1):46-56.
77. Gorkin L, Schron EB, Handshaw K, Shea S, Kinney MR, Branyon M, Campion J, Bigger JT, Sylvia SC, Duggan J, Stylianou M, Lancaster S, Ahern DK, Follic MJ. Clinical trial enrollers versus nonenrollers: The Cardiac Arrhythmia Suppression Trial (CAST)

- Recruitment and Enrollment Assessment in Clinical Trials (REACT) project. *Control Clin Trials*. 1996; 17(1):46-59.
78. Urban T, Sadow CA, Cleary JM, Harris GJ, Van Den Abbeele AD. Impact of RECIST 1.1 modifications on clinical trial eligibility. *J Clin Oncol*. 2011; 29:e16555.
79. Gordon DH, Liu Q, Kleinman B, Karrison T, Kelly M, Fleming GF. Predictors of accrual success for cooperative group trials: The Cancer and Leukemia Group B (Alliance) experience. *J Clin Oncol*. 2013; 31:TPS6645.
80. Bernardez-Pereira S, Lopes RD, Carrion MJM, Santucc EV, Soares RM, Abreu MO, Laranjeira LN, Ikeoka DT, Zazula AD, Moreira FR, Cavalcanti AB, Mesquita ET, Peterson ED, Califf RM, Berwanger O. Prevalence, characteristics, and predictors of early termination of cardiovascular clinical trials due to low recruitment: Insights from the ClinicalTrials.gov registry. *Am Heart J*. 2014; 168(2):213-9.
81. Grunfeld E, Zitzelsberger L, Coristine M, Aspelund F. Barriers and Facilitators to Enrollment in Cancer Clinical Trials. *Cancer*. 2002; 95(7):1577-83.
82. Lemieux J, Goodwin PJ, Pritchard KI, Gelmon KA, Bordeleau LJ, Duchesne T, Camden S, Speers CH. Identification of cancer care and protocol characteristics associated with recruitment in breast cancer clinical trials. *J Clin Oncol*. 2008; 26(27):4458-65.
83. Longagnani C, Bagnardi V, Zanoli L, Simoni L. A model for patients accrual in multi-site observational studies: A survival analysis approach. *Value Health*. 2011; 14:A429.
84. Rubio DM, Blank AE, Dozier A, Hites L, Gilliam VA, Hunt J, Rainwater J, Trochim WM. Developing Common Metrics for the Clinical and Translational Science Awards (CTSAs): Lessons Learned. *Clin Transl Sci*. 2015; 8(5):451-9.

85. Bennette CS, Ramsey SD, McDermott CL, Carlson JJ, Basu An, Veenstra. Predicting Low Accrual in the National Cancer Institute's Cooperative Group Clinical Trials. *J Natl Cancer Inst.* 2015; 108(2). doi: 10.1093/jnci/djv324.
86. Welton AJ, Vickers MR, Cooper JA, Meade TW, Marteau TM. Is recruitment more difficult with a placebo arm in randomized controlled trials? A quasirandomized, interview based study. *BMJ.* 1999; 318(7191):1114-7.
87. Cheng SK, Dietrich MS, Finnigan S, Dilts DM. Early indicators of accrual success: Time-to-first-patient and accrual performance at an anticipated enrollment milestone – A study of NCI-CTEP-sponsored clinical trials. *J Clin Oncol.* 2010; 28(15S):S6001.
88. Cheng SK, Dietrich MS, Dilts DM. Predicting Accrual Achievement: Monitoring Accrual Milestones of NCI-CTEP-Sponsored Clinical Trials. *Clin Cancer Res.* 2011, 17(7):1947-55.
89. Carpenter WR, Weiner BJ, Kaluzny AD, Domino ME, Lee SYD. The effects of managed care and competition on community-based clinical research. *Med Care.* 2006; 44(7):671-9.
90. Lee WR, Dignam J, Bruner D, Efstathiou JA, Yan Y, Hanks GE, Roach M, Pilepich MV, Sandler HM. Does enrollment setting influence patient attributes and outcomes in RTOG prostate cancer trials. *J Clin Oncol.* 2011; 29:S4607.
91. Carpenter WR, Fortune-Greeley AK, Zullig LL, Lee SY, Weiner BJ. Sustainability and performance of the National Cancer Institute's Community Clinical Oncology Program. *Contemp Clin Trials.* 2012; 33(1):46-54.

92. Burnett CB, Koczwara B, Pixley L, Blumenson LE, Hwang YT, Meropol NJ. Nurses' attitudes toward clinical trials at a comprehensive cancer center. *Oncol Nurs Forum*. 2001; 28(7):1187-92.
93. Califf RM, Harrington RA. American Industry and the U.S. Cardiovascular Clinical Research Enterprise [editorial]. *J Am Coll Cardiol*. 2011; 58(7):677-80.
94. Minasian LM, Carpenter WR, Weiner BJ, Anderson DE, McCaskill-Stevens W, Nelson S, Whitman C, Kelaghan J, O'Mara AM, Kaluzny AD. Translating research into evidence-based practice: the National Cancer Institute Community Clinical Oncology Program. *Cancer*. 2010; 116(19):4440-9.
95. Getz KA. Chasing veteran U.S. sites out of the enterprise: Why investigative sites are at financial risk and how it may effect sponsors and CROs [Internet]. *Applied Clinical Trials*. 2010 Nov 1 [cited 2015 Aug 24]. Accessed from: <http://www.appliedclinicaltrials.com/chasing-veteran-us-sites-out-enterprise>.
96. Gurwitz JH, Guadagnoli E, Landrum MB, Silliman RA, Wolf R, Weeks JC. The treating physician as active gatekeeper in the recruitment of research subjects. *Med Care*. 2001; 39(12):1339-44.
97. Jacobs SR, Weiner BJ, Reeve BB, Weinberger M, Minasian LM, Good MJ. Organizational and physician factors associated with patient enrollment in cancer clinical trials. *Clin Trials*. 2014; 11(5):565-75.
98. Shea S, Bigger JT, Champion J, Fleiss JL, Rolnitzky LM, Schron E, Gorkin L, Handshaw K, Kinney MR, Branyon M. Enrollment in clinical trials: Institutional factors affecting

- enrollment in the Cardiac Arrhythmia Suppression Trial (CAST). *Control Clin Trials*. 1992; 13(6):466-86.
99. Retsch-Bogart GZ, Van Dalftsens JM, Marshall BC, George C, Pilewski JM, Nelson EC, Goss CH, Ramsey BW. Highly effective Cystic Fibrosis clinical research teams: Critical success factors. *J Gen Intern Med*. 2014; 29(Supp 3):S714-S723.
 100. Butler J, Fonarow GC, O'Connor C, Adams K, Bonow RO, Cody RJ, Collins SP, Dunnmon P, Dinh W, Fiuzat M, Georgiopoulou VV, Grant S, Kim S, Kupfer S, Lefkowitz M, Mentz RJ, Misselwitz F, Pitt B, Roessig L, Schelbet E, Shah M, Solomon , Stockbridge N, Yancy C, Gheorghide M. Improving cardiovascular clinical trials conduct in the United States: Recommendation from clinicians, researchers, sponsors, and regulators. *Am Heart J*. 2015; 169(3):305-14.
 101. Taylor KM, Feldstein ML, Skeel RT, Pandya KJ, Ng P, Carbone PP. Fundamental dilemmas of the randomized clinical trial process: Results of a survey of 1,737 Eastern Cooperative Oncology Group Investigators. *J Clin Oncol*. 1994; 12(9):1796-805.
 102. Jacobs SR, Weiner BJ, Reeve BB, Weinberger M, Minasian LM, Good MJ. Organizational and physician factors associated with patient enrollment in cancer clinical trials. *Clin Trials*. 2014; 11(5):565-75.
 103. Ellis SD, Carpenter WR, Minasian LM, Weiner BJ. Effect of state-mandated insurance coverage on accrual to community cancer clinical trials. *Contemp Clin Trials*. 2012; 33(5):933-41.
 104. Morrow GR, Hickok JT, Burish TG. Behavioral aspects of clinical trials. An integrated framework from behavior theory. *Cancer*. 1994; 74(9 Suppl):2672-82.

105. Tate WR, Cranmer LD. Developing a predictive model for cancer clinical trial accrual. *J Clin Oncol.* 2014; 32(5S):6557.
106. London JW, Balestrucci L, Chatterjee D, Zhan T. Design-phase prediction of potential cancer clinical trial accrual success using a research data mart. *J Am Med Inform Assoc.* 2013; 20(e2):e260-6.
107. Gajewski BJ, Simon SD, Carlson SE. Predicting accrual in clinical trials with Bayesian posterior predictive distributions. *Stat Med.* 2008; 27(13):2328-40.
108. Moussa MAA. Planning a clinical trial with allowance for cost and patient recruitment rate. *Comput Programs Biomed.* 1984; 18(3):173-9.
109. Cantor AB. Sample size calculations for the log rank test: A Gompertz model approach. *J Clin Epidemiol.* 1992; 45(10):1131-6.
110. Lai D, Moyé, Davis BR, Brown LE, Sacks FM. Brownian motion and long-term clinical trial recruitment. *J Stat Plan Inference.* 2001; 93(1-2):239-46.
111. Carter RE. Application of stochastic processes to participant recruitment in clinical trials. *Control Clin Trials.* 2004; 25(5):429-36.
112. McNees P, Dow KH, Loerzel VW. Application of the CuSum technique to evaluate changes in recruitment strategies. *Nurs Res.* 2005; 54(6):399-405.
113. McGarvey BM, Dynes NJ, Lin BC, Anderson WH, Kremidas JP, Felli JC. A discrete event model of clinical trial enrollment at Eli Lilly and Company. In: Henderson SG, Biller B, Hsieh MH, Shortle J, Tew JD, Barton RR, editors. *Proceedings of the 2007 Winter Simulation Conference*; 2007; Washington, DC.

114. Abbas I, Rovira J, Casanovas J. Clinical trial optimization: Monte Carlo simulation Markov model for planning clinical trials recruitment. *Contemp Clin Trials*. 2007; 28(3):220-31.
115. Anisimov VV, Fedorov VV. Modelling, prediction and adaptive adjustment of recruitment in multicentre trials. *Stat Med*. 2007; 26(27):4958-75.
116. Anisimov VV. Recruitment modeling and predicting in clinical trials. *Pharm Outsourcing*. 2009; 10(1):44-8.
117. Zhang X, Long Q. Stochastic modeling and prediction for accrual in clinical trials. *Stat Med*. 2010; 29(6):649-58.
118. Zhang Q, Lai D. Fractional Brownian motion and long term clinical trial recruitment. *J Stat Plan Inference*. 2011; 141(5):1783-8.
119. Zhang X, Long Q. Joint monitoring and prediction of accrual and event times in clinical trials. *Biom J*. 2012; 54(6):735-49.
120. Mijoule G, Savy S, Savy N. Models for patients' recruitment in clinical trials and sensitivity analysis. *Stat Med*. 2012; 31(16):1655-74.
121. Gajewski BJ, Simon SD, Carlson SE. On the Existence of Constant Accrual Rates in Clinical Trials and Directions for Future Research. *Int J Stat Probab*. 2012; 1(2):43-46.
122. Tang G, Kong Y, Chang CH, Kong L, Costantino JP. Prediction of accrual closure date in multi-center clinical trials with discrete-time Poisson process models. *Pharm Stat*. 2012; 11(5):351-6.
123. Bakhsi A, Senn S, Phillips A. Some issues in predicting patient recruitment in multi-centre clinical trials. *Stat Med*. 2013; 32(30): 5458-68.

124. Deng Y, Zhang X, Long Q. Bayesian modeling and prediction of accrual in multi-regional clinical trials. *Stat Methods Med Res.* 2014 Nov 3 [Epub ahead of print].
125. Jiang Y, Simon S, Mayo MS, Gajewski BJ. Modeling and validating Bayesian accrual models on clinical data and simulations using adaptive priors. *Stat Med.* 2015; 34(4):613-29.
126. Chotai S, Porter J, Kaplan D, inventors. Method and system for modeling scenarios in clinical trials. United States provisional application 60/631,851. 2004 Nov 30.
127. Houriet JW, Horwitz AD, Subbiah A, inventors. Numoda Technologies, Inc., assignee. Automated method of graphically displaying predicted patient enrollment in a clinical trial study. United States application 13/072,334. 2011 Mar 25.
128. Goodgame JWC, Nichols GT, Brant WK, inventors. Quintiles Transnational Corp., assignee. Systems and methods for predictive analytics for site initiation and patient enrollment. United States application 13/925,262. 2013 Jun 24.
129. Zhang X, Long Q. Modeling and prediction of subject accrual and event times in clinical trials: a systematic review. *Clin Trials.* 2012; 9(6):681-8.
130. Fitts DA. Aspects of patient recruitment in clinical trials with continuous outcomes: A simulation study [dissertation]. [New Orleans]: Tulane University; 1990. 109p.
131. Kong Y. Prediction of accrual closure date in multi-center clinical trials with Poisson process models [dissertation]. [Pittsburgh]: University of Pittsburgh; 2009. 72p.
132. Anisimov VV, Fedorov VV. *Advances in Statistical Methods for the Health Sciences.* 1st ed. New York (NY): Birkhäuser Boston; c2007. Chapter 25, Design of multicentre clinical trials with random enrolment; [cited 2015 May 13]; pp387-400.

133. Anisimov VV. Discussion on the paper ‘Prediction of accrual closure date in multi-center clinical trials with discrete-time Poisson process models’, by Gong Tang, Yuan Kong, Chung-Chou Ho Chang, Lan Kong, and Joseph P. Costantino [letter]. *Pharm Stat.* 2012; 11(5):357-8.
134. Kao MCJ, Koo CC, Shem K, Wright JA, Yang L. Poster 79 Health research economics of patient recruitment: Probabilistic modeling of multicenter clinical trial recruitment for cost forecasting. *PM&R.* 2011; 3(10S1):S167.
135. Office of Cancer Centers, National Cancer Institute. P30 Cancer Center Support Grant Data Table Guide 2.0 [Internet]. Rockville, MD. National Institutes of Health/DHHS; 2015 Mar 12 [cited 2015 May 14]. 19p. Available from:
<http://cancercenters.cancer.gov/documents/CCSGDataGuide508C.pdf>.
136. Office of Cancer Centers, National Cancer Institute. NCI Website for Cancer Centers [Internet]. Rockville, MD. National Institutes of Health/DHHS; [cited 2015 Jan 25]. Available from:
<http://www.cancer.gov/researchandfunding/extramural/cancercenters/find-a-cancer-center>.
137. Clinical Trials Transformation Initiative. Comprehensive Data Dictionary 2015_March [Internet]. Durham, NC. CTTI; [cited 2015 Sep 1]. Accessible from: <http://www.ctti-clinicaltrials.org/what-we-do/analysis-dissemination/state-clinical-trials/aact-database>.
138. Harrell FE, *Regression Modeling Strategies*. 1st ed. New York: Springer 2001. 568 p.

139. Smuck B, Bettello P, Berghout K, Hanna T, Kowaleski B, Phippard L, Au D, Friel K.
Ontario Protocol Assessment Level: Clinical Trial Complexity Rating Tool for Workload
Planning in Oncology Clinical Trials. *J Oncol Pract.* 2011; 7(2):80-4.