ACCSCIENCE PUBLISHING

Journal of Clinical and Translational Research

SHORT COMMUNICATION

Prospective evaluation of the adapted Ontario Protocol Assessment Level score for predicting clinical research coordinator workload: An internal validation study

Kesley Holmes[1]*, Muhammed Idris[1], Jillian Harvey[2], Leila Forney[3], Daniel Brinton[2], Jan Morgan Billingslea[1], and Priscilla Pemu[1]

[1]Clinical Research Center, Morehouse School of Medicine, Atlanta, Georgia, United States of America

[2]Department of Healthcare Leadership and Management, College of Health Professions, Medical University of South Carolina, Charleston, South Carolina, United States of America

[3]South Carolina Clinical and Translational Research Institute, College of Medicine, Medical University of South Carolina, Charleston, South Carolina, United States of America

Abstract

Background: The escalating complexity of clinical trial protocols has considerably increased the workload for research coordinators, exacerbating staffing shortages and contributing to operational inefficiencies. These challenges are particularly pronounced at under-resourced and minority-serving research institutions, where limited capacity may hinder the implementation of trials. Early and accurate estimation of research coordinator effort is essential for effective planning, resource management, and successful clinical trial conduct. **Aim:** This study assesses the accuracy of an adopted Ontario Protocol Assessment Level (OPAL) score in predicting coordinator workload to improve operational planning in clinical research. **Methods:** A prospective observational study was conducted over a 12-month period at a Historically Black College and University medical school. Seven coordinators recorded hours for seven actively enrolling interventional trials. Estimated workloads were calculated using a published, adapted OPAL reference table, and were compared with actual hours using descriptive statistics and paired *t*-tests. To ensure consistent benchmarking, workday equivalencies (7.5 h for institutional standards and 8 h for industry standards) were applied. **Results:** There was no statistically significant difference between estimated and actual hours, with an average difference of 24.1 h (p=0.761). The mean absolute error was 167.0 h, equivalent to roughly 1 month of full-time work. **Conclusion:** The adapted OPAL score provides a practical tool for estimating coordinator workload and aligning staffing with protocol complexity, including in under-resourced settings. However, broader multi-site validation is required to confirm its generalizability and to support its integration into feasibility planning. **Relevance for patients**: Accurate workload forecasting enhances trial efficiency, supporting timely, high-quality studies, and accelerating access to new treatments.

Keywords: Workload estimation; Ontario Protocol Assessment Level score; Clinical trial operations; Research coordinator workload; Protocol complexity; Implementation science; Workforce planning; Coordinator staffing models

***Corresponding author:**
Kesley Holmes
(ktyson@msm.edu)

Citation: Holmes K, Idris M, Harvey J, *et al.* Prospective evaluation of the adapted Ontario Protocol Assessment Level score for predicting clinical research coordinator workload: An internal validation study. *J Clin Transl Res*. 2025;11(5):106-112.
doi: 10.36922/JCTR025260032

Received: June 28, 2025

Revised: August 12, 2025

Accepted: August 12, 2025

Published online: August 25, 2025

1. Introduction

The increasing complexity of clinical trial protocols has significantly amplified the demands placed on research coordinators, who serve as the operational backbone of study implementation. These professionals are responsible for a wide range of critical tasks, including regulatory compliance, patient engagement, data collection, adverse event reporting, and visit schedule adherence, all of which have become more time-consuming and resource-intensive. As protocols become more intricate, workload imbalances among coordinators are becoming increasingly common, contributing to elevated stress, burnout, and staff turnover.[1-5] These challenges are further exacerbated by staffing shortages and funding limitations at many academic and community-based research sites.

Accurately estimating research coordinator effort is essential for informed decision-making around staffing, resource allocation, and study feasibility. Effective planning depends on the ability to forecast the operational and administrative complexity of a study before it begins.[6,7] Without reliable workload prediction models, sites risk under- or over-allocating personnel, potentially compromising compliance and performance.

The Ontario Protocol Assessment Level (OPAL) score was previously developed to quantify protocol complexity by assigning numerical values to objective trial characteristics such as intervention type, number of study procedures, and frequency of patient visits.[8] While the OPAL score has gained broad adoption as a baseline tool, it has limitations when used in isolation. Specifically, it does not account for site-level operational variables that meaningfully influence workload, such as brief recruitment windows, complex specimen handling requirements, language barriers, or high-intensity data queries and monitoring activities. To address these limitations, it is recommended that the score be adapted by reweighing existing elements and incorporating additional workload drivers.[8,9]

Tyson *et al.*[9] developed an adapted OPAL score that integrates supplemental complexity indicators and links the score to observed coordinator effort using retrospective data from a diverse portfolio of clinical trials. Their analysis revealed a strong linear relationship between the adapted OPAL score and actual hours worked by coordinators ($\beta = 77.22$; $p=0.01$; $R^2 = 0.78$), resulting in a practical reference table for estimating staff effort during trial planning. However, this tool has not yet been validated.

Validation of workload estimation tools is critical in confirming their utility, accuracy, and generalizability across research settings. A validated tool provides measurable confidence that its estimates reliably reflect the construct being assessed—in this case, the Clinical Research Coordinator (CRC) effort. According to Streiner *et al.*,[10] validation requires demonstrating that tool-derived predictions align with observed outcomes, ideally across multiple contexts and study designs. For tools predicting operational metrics, prospective validation enhances external validity by assessing real-time workflows rather than relying solely on retrospective analyses.[11,12]

In this study, prospective validation methods are applied to evaluate the adapted OPAL score by comparing its predicted coordinator workload against actual hours logged across seven interventional trials at a single site. This approach aligns with best practices in validating prediction models and workload estimation frameworks.[13-15]

2. Materials and methods

This prospective observational study was conducted at an academic clinical research site to evaluate the accuracy of the adapted OPAL score in predicting research coordinator workload. Between January 01 and December 31, 2024, seven CRCs tracked the hours they spent managing seven actively enrolling interventional trials. The selected studies varied in sponsor type (industry-sponsored vs. federally funded) and intervention type (drug vs. behavioral).

The adapted OPAL score was calculated for each trial based on predefined criteria, including procedural volume, visit intensity, monitoring requirements, and biospecimen complexity. Estimated coordinator workload hours were then derived using a previously published reference table developed by Tyson *et al.*,[9] which was constructed from retrospective coordinator time-tracking data collected across a range of experience levels. This approach was intended to produce weights representative of overall coordinator effort rather than a single experience stratum. The reference table maps OPAL score tiers to predicted effort (Table 1).

Each CRC prospectively logged actual hours worked per protocol using a standardized digital time-tracking system. Data were reconciled weekly to ensure completeness and accuracy. Estimated workload hours were also converted into workday equivalents using both a 7.5-h academic standard and an 8-h industry standard. This conversion was performed solely to facilitate cross-sector benchmarking. Such conversions are common in resource planning models, allowing institutions to interpret workload estimates in the context of their operational norms.

2.1. Statistical analysis

To assess the agreement between the estimated and actual workload, descriptive statistics–including mean absolute

error (MAE) and mean difference–were calculated, consistent with early-stage predictive model validation practices.[16] A paired Student's *t*-test was used to assess differences between estimated and actual hours, as both values were generated from the same coordinator-study pairing. Unpaired Student's *t*-tests were used for subgroup analyses (i.e., sponsor type and intervention type) to compare mean values between independent groups.[17]

These methods enable both absolute and relative evaluations of prediction accuracy, highlighting areas for potential refinements of the adapted OPAL score. Workday equivalencies were computed using 7.5-h institutional and 8-h industry standards. This adjustment ensures consistency when comparing internal workloads to external benchmarks. *p*-values were not calculated for individual trials because each trial's estimated and actual values represent a single paired observation for the entire 12-month period, making statistical significance testing at the trial level mathematically inappropriate. The reported *p*-value for the "estimated versus actual comparison" was calculated using the aggregated paired dataset across all trials, allowing for appropriate variance estimation.

This study did not require Institutional Review Board approval, as it was classified as a quality improvement and operational research initiative aimed at enhancing internal clinical trial management processes. No identifiable private information was collected, and the project was not considered human subjects research.

3. Results

The seven interventional trials included in this study consisted of five Phase 3 trials, one Phase 2/3 hybrid trial, and one Phase 2 trial, spanning a range of therapeutic areas.

Table 1. Reference table for the adapted Ontario Protocol Assessment Level score

Adapted Ontario Protocol Assessment Level score	Estimated hours (6-month period)	Estimated hours per month
5.5	30.7	5.1
6.0	69.3	11.5
6.5	107.9	18.0
7.0	146.5	24.4
7.5	185.1	30.9
8.0	223.7	37.3
8.5	262.3	43.7
9.0	301.0	50.2
9.5	339.6	56.6

Note: Estimated coordinator workload hours were calculated using the adapted Ontario Protocol Assessment Level reference table previously published by Tyson *et al.*[9]

The MAE between the adapted OPAL-based estimated workload and the actual coordinator hours logged for the 12-month study period was 167.0 h, equivalent to approximately 22.3 workdays (4.5 weeks or 1.0 month) using a 7.5-h institutional workday and 20.9 workdays (4.2 weeks or 1.0 month) using an 8-h industry-standard workday. Despite this variability, the average difference between estimated and actual hours across all trials was relatively modest at 24.1 h and was not statistically significant (t = 0.32, p=0.761). This difference represents approximately 7–8% of a full-time coordinator's annual effort. A detailed summary is provided in Table 2 and Figure 1.

When analyzed by sponsor type, industry-sponsored trials required more coordinator time, with an average of 422.8 actual hours compared to 246.0 h for federally funded trials. This difference indicated a trend toward significance (t = −2.06, p=0.095), suggesting that sponsor type may be influential in predicting coordinator burden. Although industry trials demonstrated slightly higher average adapted OPAL scores, the difference was not statistically significant (Figure 2). No substantial differences between drug and behavioral intervention trials were observed

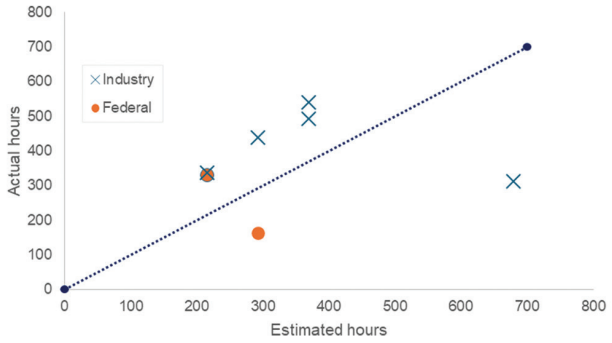

Figure 1. Estimated versus actual hours over a 12-month period

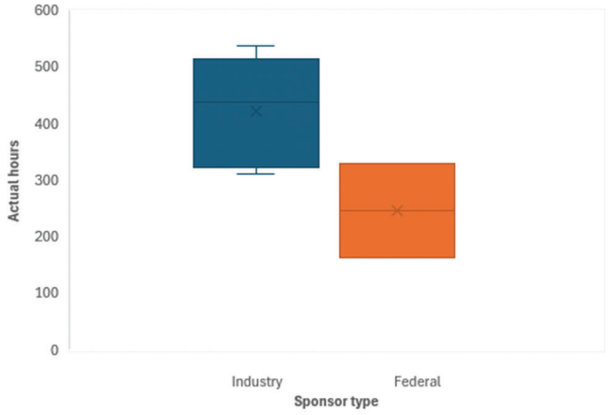

Figure 2. Comparison of actual coordinator hours by sponsor type

Table 2. Summary of study characteristics and results by protocol

Trial number	Adapted OPAL score	Trial phase	Sponsor type	Intervention	Estimated hours[a]	Actual hours	Difference
1	7.5	3	Industry	Drug	370.2	538	167.8
2	6.5	3	Industry	Drug	370.2	492	121.8
3	7.0	2/3	Industry	Drug	293.0	438	145.0
4	9.5	3	Industry	Drug	679.2	310	−369.2
5	6.5	3	Federal	Behavioral	215.8	330	114.2
6	6.5	3	Industry	Drug	215.8	336	120.2
7	7.0	2	Federal	Drug	293.0	162	−131.0

Note: [a]Hours were estimated for 12 months.
Abbreviation: OPAL: Ontario Protocol Assessment Level.

in estimated or actual coordinator hours. However, interpretation of these comparisons is limited by the small number of trials and the limited behavioral studies in the dataset.

These findings suggest that the adapted OPAL score is a promising tool for estimating workload, particularly in later-phase trials characterized by diverse and complex operational requirements. No statistically significant difference was observed between the estimated and actual hours, which may be due to the small sample size. Because the MAE reflects a 12-month study period, even if the difference was robust in larger samples, it would remain relatively small, representing merely 7–8% of the annual CRC workload. Variability in prediction accuracy across trials suggests the tool may benefit from further refinement to account for study-specific factors.

4. Discussion

This study builds on the foundational work of Tyson *et al.*,[9] who introduced the adapted OPAL score as a tool for estimating CRC effort. Unlike the original retrospective analysis, the present study utilizes the tool prospectively, demonstrating its real-time utility for operational planning and staffing allocation. Future multi-center validation studies should consider stratifying results by coordinator experience or including experience as a covariate in predictive modeling to further enhance predictive accuracy.

The findings indicate that the adapted OPAL score can predict coordinator effort with a high degree of accuracy. No statistically significant difference was observed between the estimated and actual hours, and the MAE reflected a manageable variance of approximately 8%, equivalent to less than one month of full-time work. Although absolute differences in hours were sometimes large for individual studies, the percentage variance was modest, thereby supporting the adapted OPAL's utility for budget and staffing estimates. Future multi-site, prospective time-tracking could be used to refine the weighting system further and reduce variability, particularly for protocols with high operational complexity or atypical team structures. These findings highlight the value of the tool in helping research sites anticipate and manage staffing needs, potentially reducing the risk of understaffing, missed milestones, and staff burnout. As pressure mounts for research operations to become more efficient, particularly in light of proposed reductions in administrative cost allowances for grants, accurate workload estimation and budgeting will become increasingly critical for maintaining operational sustainability.

By providing workload estimates in both 7.5-h (academic) and 8-h (industry) workday equivalents, the score enhances its practical relevance for a broad range of stakeholders, including academic health centers, contract research organizations, and community-based research institutions. This flexibility ensures that operational estimates remain meaningful regardless of institutional norms, improving cross-site comparability and planning.

By quantifying protocol complexity and translating it into time-based estimates, the adapted OPAL score addresses a key limitation in traditional trial feasibility practices, which often rely on subjective judgment or historical precedent.[7,18,19] The adapted OPAL score allows for a more data-driven and scalable approach to workload forecasting and supports a more efficient staffing model for better resource alignment and enhanced financial sustainability. The adapted OPAL builds on a methodology designed to include both protocol-driven and ancillary activities; however, some unstructured tasks will inevitably remain unmeasured. Pairing OPAL estimates with periodic portfolio-level reviews and integration into a Clinical Trial Management System (CTMS) can help identify emerging or unaccounted workload, enabling mid-course staffing adjustments.

Further tool development could include integration with CTMS, allowing real-time updates to workload projections in response to evolving study demands. This dynamic approach would improve operational agility and allow research teams to respond proactively to mid-study shifts, such as protocol amendments and accelerated recruitment timelines.

In addition, the observed trend toward higher coordinator burden in industry-sponsored trials, though not statistically significant, aligns with findings from prior studies that suggest increased operational complexity in commercially funded studies.[20] These trials often include more rigorous documentation requirements, frequent monitoring visits, and a higher frequency of protocol amendments, all of which demand greater coordinator time. These findings highlight the importance of adapting workload estimations to sponsor characteristics during site-level planning.

Several limitations remain in this study, including the small sample size and single-site design, which limit generalizability. The limited number of federally funded trials in the sample (n = 2, one of which was behavioral) created an imbalance in sponsor representation, constraining the ability to assess whether the adapted OPAL tool's predictive accuracy differs meaningfully between sponsor types. Future multi-site studies across diverse trial portfolios are needed to further validate the tool and assess its impact on trial performance metrics. With broader validation, the adapted OPAL score could be integrated into feasibility review workflows, budget justification tools, and institutional staffing frameworks. Future research should explore whether improved workload forecasting correlates with enhanced study outcomes, including faster recruitment, fewer protocol deviations, and improved data quality. A multi-site evaluation would strengthen generalizability and enable the development of benchmarking tools to compare coordinator efforts across institutions.

Importantly, this study was conducted at a medical school of a historically Black College and University—a community-based, minority-serving, and under-resourced institution. As such, it provides critical insight into the operational realities of underrepresented research sites. These institutions often carry a disproportionate operational burden and systemic barriers to trial participation and sustainability.[21,22] The successful application of the adapted OPAL score in this context highlights its potential as an equitable, scalable tool for supporting workload planning and staffing decisions. These adaptations were intentionally designed to address the disproportionate operational burden and systemic barriers faced by similar sites, and future iterations could incorporate additional site-specific modifiers for broader applicability. Broader implementation of such models may help reduce disparities in site performance, build long-term research capacity, and promote workforce sustainability, particularly in settings vital to expanding clinical research access to underserved populations.

5. Conclusion

When applied at the outset of a clinical trial, the adapted OPAL score offers a reliable, evidence-based method for forecasting coordinator workload and aligning staffing needs with protocol complexity. By translating study requirements into projected full-time equivalent allocations, the tool supports more informed feasibility assessments, facilitates sponsor-site negotiations, and improves operational readiness. Early estimation of effort also enables proactive staffing and budget forecasting, which are critical elements of efficient and sustainable trial execution.

Importantly, no statistically significant difference was observed between the estimated and actual hours, thereby supporting the adapted OPAL score's accuracy. The MAE of 167.0 h (approximately 1 month of full-time work) provides a practical benchmark for staffing calibration based on institutional norms. However, these results should be interpreted in the context of the study's limited sample size and single-site design. Further validation across institutions, trial phases, and therapeutic areas is needed to strengthen the generalizability and support integration of the tool into feasibility planning workflows.

This study was conducted at a medical school of a historically Black College and University, a community-based, minority-serving, and under-resourced institution. The successful application of the adapted OPAL score in this setting underscores its practicality and relevance for research sites facing systemic barriers and operational constraints. Broader adoption of this tool could help reduce disparities in trial performance, improve infrastructure, and promote workforce sustainability in underserved settings.

With continued refinement and integration into CTMS, the adapted OPAL score could serve as a standard tool for feasibility reviews, budget planning, and staffing models. Its application may ultimately enhance trial efficiency, support research staff, and promote equity in clinical research.

Acknowledgments

The authors thank the research coordinators and administrative staff whose contributions were vital to this study.

Funding

This project was supported in part by the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Numbers UL1TR002378, UL1 TR001450, and UM1 TR005294, and by the National Institute on Minority Health and Health Disparities under Award Number 1U24MD015970. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Kesley Holmes
Formal analysis: Kesley Holmes, Muhammed Idris, Daniel Brinton
Investigation: Kesley Holmes, Jan Morgan Billingslea
Methodology: Kesley Holmes, Muhammed Idris
Writing–original draft: All authors
Writing–review & editing: All authors

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

Data are available from the corresponding author on reasonable request.

Further disclosure

A portion of the findings included in this manuscript was previously presented as a poster at the Southeast Clinical and Translational Science Alliance Conference held in Pine Mountain, Georgia, United States, in March 2025. The presentation, titled "A validation study of the adapted OPAL workload estimation tool," shared preliminary results from the study to facilitate scholarly discussion and obtain feedback. The content has since been expanded and refined for this manuscript submission. The poster has not been published or released by the conference organizers. In addition, this paper was polished and edited with the assistance of artificial intelligence-powered tools, specifically ChatGPT and Grammarly. These tools were used solely to enhance clarity, grammar, and formatting. All content, analyses, and ideas are original and authored by the researchers; the project was not generated or conceptualized by artificial intelligence.

References

1. Getz KA, Campo RA. New benchmarks characterizing growth in protocol design complexity. *Ther Innov Regul Sci*. 2018;52(1):22-28.

   doi: 10.1177/2168479017713039

2. Getz KA, Wenger J, Campo RA, Seguine ES, Kaitin KI. Assessing the impact of protocol design changes on clinical trial performance. *Am J Ther*. 2008;15(5):450-457.

   doi: 10.1097/MJT.0b013e31816b9027

3. Florence Healthcare. *Clinical Trial Complexity Index Report*; 2023. Available from: https://florencehc.com/resources/clinical-trial-complexity-index [Last accessed on 2025 Apr 20].

4. Florence Healthcare. *Complexity Costs: Hidden Tolls in Clinical Trial Operations*; 2023. https://florencehc.com/resources/complexity-costs [Last accessed on 2025 Apr 20].

5. Gwede CK, Johnsson DJ, Roberts C, Cantor AB. Burnout in clinical research coordinators in the United States. *Oncol Nurs Forum*. 2005;32(6):1123-1130.

   doi: 10.1188/05.onf.1123-1130

6. Speicher LA, Fromell G, Avery S, *et al.* The critical need for academic health centers to assess the training, support, and career development requirements of clinical research coordinators: Recommendations from the clinical and translational science award research coordinator taskforce. *Clin Transl Sci*. 2012;5(6):470-475.

   doi: 10.1111/j.1752-8062.2012.00423.x

7. Morin DJ. Harmonizing protocol complexity with resource management and capacity planning at clinical research sites. *Ther Innov Regul Sci*. 2020;54:978-987.

   doi: 10.1007/s43441-020-00120-8

8. Smuck B, Bettello P, Berghout K, *et al.* Ontario protocol assessment level: Clinical trial complexity rating tool for workload planning in oncology clinical trials. *J Oncol Pract*. 2011;7(2):80-84.

   doi: 10.1200/JOP.2010.000051

9. Tyson K, Harvey J, Forney L, Brinton D. Resource management and capacity planning for clinical trial sites. *J Clin Transl Res.* 2024;10(4):229-236.

   doi: 10.36922/jctr.24.00022

10. Streiner DL, Norman GR, Cairney J. *Health Measurement Scales: A Practical Guide to their Development and Use*. 5th ed. Oxford: Oxford Academic; 2015.

    doi: 10.1093/med/9780199685219.001.0001

11. Kottner J, Audigé L, Brorson S, *et al.* Guidelines for Reporting reliability and agreement studies (GRRAS). *J Clin*

Epidemiol. 2011;64(1):96-106.

doi: 10.1016/j.jclinepi.2010.03.002

12. Wu K, Wu E, DAndrea M, *et al.* Machine learning prediction of clinical trial operational efficiency. *AAPS J*. 2022;24(3):57.

doi: 10.1208/s12248-022-00703-3

13. Moons KG, Altman DG, Reitsma JB, *et al.* Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. *Ann Intern Med*. 2015;162(1):W1-W73.

doi: 10.7326/M14-0698

14. Kumar B, Verma A, Verma P. Optimizing resource allocation using proactive scaling with predictive models and custom resources. *Comput Electr Eng*. 2024;118:109419.

doi: 10.1016/j.compeleceng.2024.109419

15. Qureshi SM, Purdy N, Neumann WP. Developing a modelling approach to quantify quality of care and nurse workload - field validation study. *Oper Res Health Care*. 2021;29:100301.

doi: 10.1016/j.orhc.2021.100301

16. Eichenseer P, Hans L, Winkler H. A data-driven machine learning model for forecasting delivery positions in logistics for workforce planning. *Supply Chain Anal*. 2025;9:100099.

doi: 10.1016/j.sca.2024.100099

17. Kim TK. T test as a parametric statistic. *Korean J Anesthesiol*. 2015;68(6):540-546.

doi: 10.4097/kjae.2015.68.6.540

18. Good MJ, Lubejko B, Humphries K, Medders A. Measuring clinical trial-associated workload in a community clinical oncology program. *J Oncol Pract*. 2013;9(4):211-215.

doi: 10.1200/jop.2012.000797

19. Richie A, Gamble D, Tavlarides A, Strok K, Griffin C. Establishing the link between trial complexity and coordinator capacity. *Clin Res*. 2020;34:8-16.

20. WCG. *2024 Clinical Research Site Challenges Report*. WCG; 2024. Available from: https://www.wcgclinical.com/wp-content/uploads/2024/10/wcg_2024_clinical_research_site_challenges_report.pdf?utm_source=chatgpt.com [Last accessed on 2025 Apr 04].

21. Loree JM, Anand S, Dasari A, *et al.* Disparity of race reporting and representation in clinical trials leading to cancer drug approvals from 2008 to 2018. *JAMA Oncol*. 2019;5(10):e191870.

doi: 10.1001/jamaoncol.2019.1870

22. Chen MS Jr, Lara PN, Dang JH, Paterniti DA, Kelly K. Twenty years post-NIH Revitalization Act: Enhancing minority participation in clinical trials (EMPaCT): Laying the groundwork for improving minority clinical trial accrual: Renewing the case for enhancing minority participation in cancer clinical trials. *Cancer*. 2014;120(Suppl 7):1091-1096.

doi: 10.1002/cncr.28575