



# Insights from a Multi-company Workshop to Apply a Patient Participation Burden Algorithm to Protocol Data

Zachary Smith, MA<sup>1</sup> · Emily Botto, BA<sup>1</sup> · Christopher Carney, BS<sup>2</sup> · Abhishek Bagga, MBA<sup>2</sup> · Bazgha Qutab, MBA<sup>2</sup> · Kenneth Getz, MBA<sup>1</sup>

Received: 25 May 2022 / Accepted: 30 September 2022  
© The Author(s), under exclusive licence to The Drug Information Association, Inc 2022

## Abstract

**Background** Utilizing a participation burden algorithm developed in a previous study, Tufts CSDD, in collaboration with ZS, led a workshop among 8 pharmaceutical companies to validate the methodology of benchmarking the participation burden of a set of retrospective protocols and comparing these data to a prospective protocol design.

**Methods** Eight participating companies collected data for 66 retrospective protocols and participation burden scores were calculated for each. Data from one prospective protocol was provided and prospective burden scores were compared to mean retrospective protocol burden for each company. Participating companies provided feedback on data collection process and final reports.

**Results** Comparisons between retrospective and prospective burden scores revealed higher comparative burden in lab and blood procedures. Companies were able to gather most requested data, but some variables hypothesized to affect burden were not available to sponsors. Time constraints were reported as a challenge throughout the data collection process.

**Conclusions** Feedback indicated the need for establishing a larger database to enable comparisons between protocols with the same therapeutic area and indication. Investigating the impact of standard of care burden by indication on overall participation burden and encouraging sponsors to collect more accurate data contributing to participation burden at the site level are also important takeaways from this exercise.

**Keywords** Patient burden · Protocol design · Patient engagement · Patient centricity

## Introduction

Over the past several years, the Tufts Center for the Study of Drug Development (Tufts CSDD) and ZS—a global management consulting and professional services firm—have collaborated to empirically measure the subjective burden that patients experience as participants in clinical trials. The primary objective of this collaborative research is to develop a systematic approach that protocol design authors and clinical teams can use to anticipate and mitigate patient participation burden prior to protocol finalization.

In an initial study, a survey was conducted in 2019 among 591 US-based patients who rated the perceived burden of

60 distinct, commonly performed, clinical trial procedures. From these ratings, Tufts CSDD and ZS calculated average burden scores for individual procedures—by select participant demographic and disease subgroups—and derived an algorithm quantifying overall participation burden in the clinical trial. The algorithm was then tested against a convenience sample of historical protocol designs and their corresponding clinical trial performance outcomes (e.g., cycle times, recruitment and retention rates). The study results showed that the burden score calculated by the algorithm was significantly correlated with screen failure rates and treatment duration (i.e., first patient first visit to last patient last visit) [1].

A year later, a second survey was conducted among 2,680 global patients. In addition to gathering ratings of perceived procedural burden, the survey also gathered patient perceptions of non-procedural burden (such as visit durations and location where the procedure was conducted) and lifestyle restrictions imposed by trial participation (i.e., travel, diet,

✉ Zachary Smith  
Zachary.Smith605922@tufts.edu

<sup>1</sup> Tufts CSDD, 145 Harrison Ave., Boston, MA 02111, USA

<sup>2</sup> Evanston, USA

exercise restrictions). When non-procedural burden was accounted for, the enhanced algorithm was highly associated with, and predictive of, study conduct and total clinical trial duration (i.e., protocol approval to database lock), screen failure rate, and the number of substantial protocol amendments [2].

Although the results of the 2019 and 2020 studies offered valuable insights, they were both limited to convenience samples of retrospective protocols. As a critical next step, Tufts CSDD and ZS turned their attention to the prospective application of the algorithm. The purpose of this new study was to give sponsor companies hands-on experience with the algorithm and to determine the feasibility of applying the participant burden algorithm in the actual protocol design drafting stage.

Tufts CSDD and ZS used a workshop approach for this new study. Senior clinical operations and data science representatives from several pharmaceutical companies provided data on previously completed protocols and on prospective protocols that had not yet been completed. The retrospective protocols in the workshop gave participating companies experience with the data gathering and reporting methodology established in the earlier studies. Tufts CSDD and ZS used each company's retrospective data to prepare comparative benchmarks. Data from prospective protocols were scored and compared to these benchmarks. Tufts CSDD and ZS prepared custom reports and discussed the results with each company participating in the workshop.

## Methodology

### Data Collection

Mid-year 2021, Tufts CSDD and ZS met with 8 companies in a virtual workshop to review and discuss past participant burden algorithm studies and the process for collecting data for the new study. Workshop participants reviewed the data collection worksheet used in the past Tufts CSDD-ZS studies. Participating companies next provided feedback on which variables could be easily produced, which would be challenging to provide, and which were infeasible. Company feedback was incorporated, and the worksheet was revised until a consensus had been reached. Workshop attendees then completed the retrospective worksheet for their selected protocols. Table 1 provides an outline of the worksheet and the data collected.

Each company gathered data on 7–9 retrospective protocols and a single prospective protocol. The same worksheet was used to collect retrospective and prospective data, with 1 modification for the prospective worksheet: procedure counts were provided by visit instead of total counts for the

**Table 1** Data collection worksheet description

	Number of questions
Protocol data	
Trial characteristics	6
Lifestyle restrictions	8
Caregiver requirements	5
Patient dropout	5
Participant data	
Participant subgroups	6
Participant characteristics	2
Procedure type	
Medication	6
Lab and blood	9
Routine procedure	10
Non-invasive procedure	5
Invasive procedure	11
Imaging	8
Questionnaire	4
Other	5

trial overall. This would allow for analysis of individual trial visits to potentially identify unusually burdensome visits.

After completing the retrospective data collection, companies began collecting the prospective data, and Tufts CSDD began scoring the retrospective protocols according to the previously developed algorithm.

### Analysis and Reporting

Procedural burden was calculated for each protocol based on the procedural burden scores provided in the survey from stage 2. These procedural burden scores accounted for demographic subgroup data provided by the companies, as well as the therapeutic area of the target disease. The procedural burden was then adjusted according to the various protocol characteristics estimated to have inflation or deflation effects on the perceived burden of participation as discussed in the manuscript by Smith et al. [2]. Once prospective data collection was completed, the prospective scores were calculated the same way.

A report format outlined the burden scores of the overall sample, compared individual company protocols to the rest of the sample, and compared each company's prospective protocol to the company's own retrospective sample. Initially the report presented mean scores, however given the wide range of scores, as well as outliers, median scores provided a better representation of the sample. The report presented a description of the sample by phase and therapeutic area. It also presented median total burden scores and procedural burden scores by procedure type. In recognition of the high

relative complexity and burden associated with oncology protocols, benchmarks were also prepared for oncology vs non-oncology protocols. For each prospective protocol, procedure counts and burden scores of unique procedures were calculated. The mean procedure count as well as the mean burden of all unique procedures were calculated. Procedures that were performed more than the mean number of times for the protocol were considered ‘high frequency’ procedures; procedures with perceived burden above the mean for the protocol were considered ‘high burden.’ The report also presented a figure illustrating the relative procedural burden contribution of procedures in the protocol—those that were ‘high burden, high frequency;’ those that were ‘high burden, low frequency;’ those that were ‘low burden, high frequency;’ and those that were ‘low burden, low frequency.’

**Table 2** Data characteristics

Protocol characteristic	Percent of protocols in sample ( <i>n</i> = 66) (%)
Phase	
Phase II	45.5
Phase III	54.6
Therapeutic area	
Cardiology	4.6
Dermatology	1.5
Endocrinology	7.6
Immunology	21.2
Infectious disease	10.6
Neurology	4.6
Oncology	40.9
Respiratory	4.6
Rheumatology	4.6

**Table 3** Total burden and procedural burden by burden type

Burden type	<i>n</i>	Mean (CoV)	Median	Range
Overall				
Total burden	66	10,534.4 (1.0)	8740.1	842.4–57,570.8
By procedure type				
Medication burden	66	752.6 (1.7)	350.1	23.9–7785.2
Lab and blood burden	66	1910.0 (1.1)	1455.1	89.3–16,108.1
Routine procedure burden	66	982.9 (0.8)	717.4	0.0–3693.2
Non-invasive procedure burden	66	183.1 (1.2)	100.6	0.0–1147.1
Invasive procedure burden	66	84.0 (2.5)	0.0	0.0–1208.3
Imaging procedure burden	66	316.8 (1.8)	85.8	0.0–3600.9
Questionnaire procedure burden	66	858.6 (1.3)	519.5	0.0–5539.4
Other procedure burden	66	1475.5 (4.5)	0.0	0.0–45,209.7

## Results

The dataset was disproportionately weighted toward oncology trials, with 41% of the set representing this therapeutic area. Analyses were conducted on overall protocols as well as oncology and non-oncology subgroups. Immunology also made up a considerable proportion of the provided protocols, while the remaining protocols were distributed among seven other therapeutic areas. The sample was almost evenly split between Phase II and III trials, with a slightly higher percentage of Phase III (Table 2).

Procedures were broken out by types, with the procedure type ‘lab and blood’ making up the highest percentage of overall burden. The frequency of blood draws in most protocols, in addition to the high burden associated with blood draws, contributed to this higher percentage of the mean burden score in comparison to other procedure types. Although invasive procedures carried a high score per individual procedure, their low frequency led to their small role in driving burden. It’s also worth noting that despite the fact that invasive procedures were more common among oncology trials than non-oncology trials, their contribution to total burden remained relatively small. Routine procedure scores also represented a large percentage of the average overall burden and were heavily dependent on physical exams. Although an individual physical exam does not carry a large burden score within the algorithm, its frequency was such that a low initial burden added up to make it one of the most burdensome procedures in the set (Table 3).

Lab and blood procedures’ share of overall burden increased in the oncology sample. Imaging also carried a higher proportion of burden for oncology protocols, while routine procedures made up a smaller portion of the overall burden. Perhaps unsurprisingly, oncology protocols also presented higher overall median burden scores as well as higher scores for each procedure type except for questionnaire burden, when compared to non-oncology trials. Oncology

procedure type median scores showed particularly high relative burden to non-oncology trials in medication, lab and blood, invasive procedures (e.g., biopsies, catheterizations, etc.), and imaging (Table 4).

Procedural data collection proceeded efficiently, with few complications and little need for clarification. Most procedures for the protocols assessed fell within the 58 common trial procedures with burden estimates, which were provided within the data collection worksheet. Some additional procedures fell outside of these 58 defined procedures; in these cases, the company representatives were able to categorize the outlying procedure into one of the procedure types, and a burden score was calculated based on the categorization. Some examples of the few procedures that required proxy categorization included molecular cytogenetics and flow cytometry.

Some barriers to providing procedure frequencies arose when a single procedure could be considered two different procedure types: for example, a clinician conducting a verbal questionnaire during a physical exam. Reporting this as two distinct procedures, both a physical exam and a clinic questionnaire, could result in a burden score that is higher than the burden actually perceived by the participant.

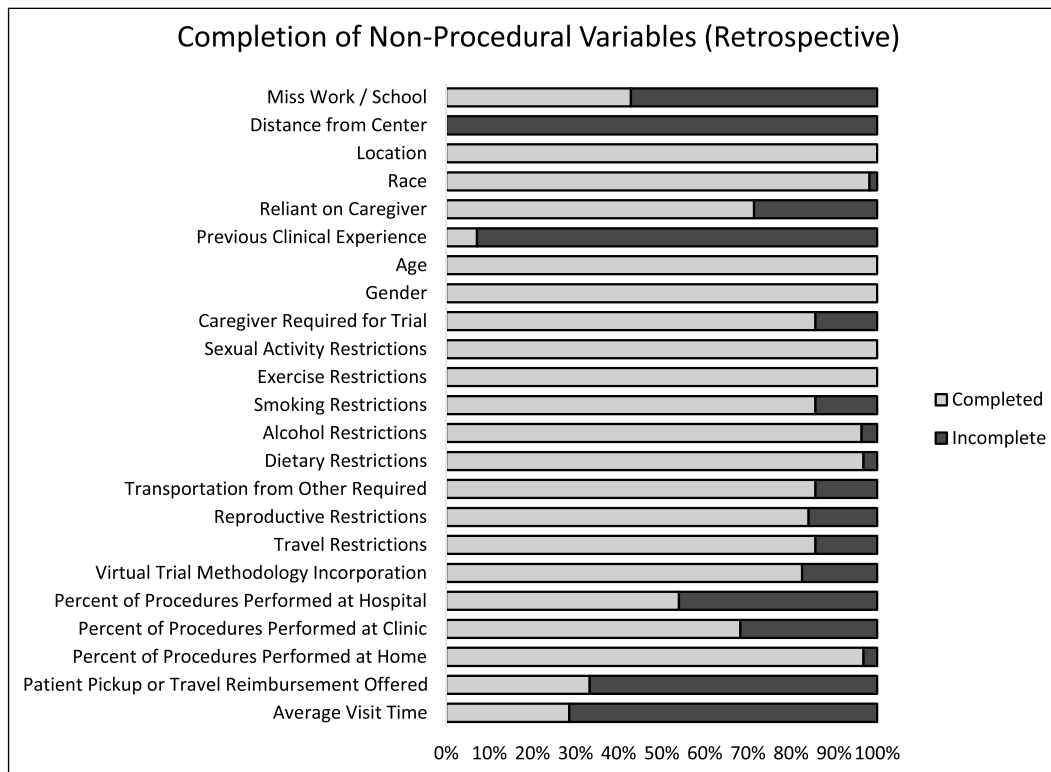
Procedure counts for blood samples, additionally, harbored some nuances for calculating perceived burden. Protocols often cite the number of tests performed on blood samples as the procedure frequency. However, if the tests are performed using the same blood sample, participants will only experience the burden of a single blood

draw. The tests performed on that blood sample, regardless of the amount, should not affect participant burden. Therefore, participating companies in the workshop were instructed to only count blood draws for blood sample procedure frequency. According to representatives, there were barriers to acquiring the number of blood draws, and the numbers reported were not always accurate.

Companies reported significant challenges in acquiring certain non-procedural data, particularly average visit time, participant pickup or travel reimbursement offered, previous clinical trial experience, distance from center, and whether the clinical trial protocol required the participant to miss work or school. According to companies in the workshop, although data on participant's previous clinical trial experience are not directly collected, various proxies are often available, such as eligibility criteria that require the participant to have been part of a previous study of that treatment or a similar treatment. Exclusion criteria may also provide insight on this variable through requirements that prevent participants from having been enrolled in any trials within a specified timeframe. The remaining variables—distance from center, average visit time, participant pickup or travel reimbursement offered, and whether the clinical trial protocol required the participant to miss work or school—were the most difficult to collect, and less than 50% of participating companies in the workshop were able to provide these variables for any protocol. A visualization of non-procedural data collection completion rates is presented in Fig. 1.

**Table 4** Total burden and procedural burden by burden type—oncology vs. non-oncology

Burden type	<i>n</i>	Mean (CoV)	Median	Range
<b>Oncology</b>				
Total burden	27	11,380.2 (0.8)	9527.4	1214.0–40,011.7
Medication burden	27	804.6 (1.0)	591.9	38.0–2980.4
Lab and blood burden	27	2504.9 (1.2)	2068.5	429.7–16,108.1
Routine procedure burden	27	989.9 (0.7)	725.7	53.6–3241.7
Non-invasive procedure burden	27	189.6 (1.1)	100.6	0.0–743.4
Invasive procedure burden	27	105.8 (1.6)	65.3	0.0–849.3
Imaging procedure burden	27	640.1 (1.1)	486.6	0.0–3600.9
Questionnaire procedure burden	27	855.6 (1.1)	516.5	0.0–4270.0
Other procedure burden	27	851.5 (5.1)	0.0	0.0–22,479.7
<b>Non-oncology</b>				
Total burden	39	9948.8 (1.1)	7544.6	842.4–57,570.8
Medication burden	39	716.6 (2.2)	147.4	23.9–7785.2
Lab and blood burden	39	1498.2 (0.6)	1330.8	89.3–3941.8
Routine procedure burden	39	978.0 (0.9)	710.1	0.0–3693.2
Non-invasive procedure burden	39	178.7 (1.3)	83.8	0.0–1147.1
Invasive procedure burden	39	68.9 (3.4)	0.0	0.0–1208.3
Imaging procedure burden	39	92.9 (2.9)	0.0	0.0–1579.1
Questionnaire procedure burden	39	860.6 (1.4)	522.5	0.0–5539.4
Other procedure burden	39	1907.6 (4.1)	0.0	0.0–45,209.7



**Fig. 1** Completion of non-procedural burden variables

In addition to barriers in access to certain procedural and non-procedural datapoints, participating companies reported challenges in availability of employees to conduct data collection. Many participating companies reported difficulties managing what was reported as a time-consuming data collection process. Staff dedicated to data collection may be an important factor in effective and accurate data.

In a few instances, representatives were not able to complete data collection due to personal workload. In these cases, the research team performed data collection based on the full protocol, provided by the given company. For the retrospective data collection, only one company chose this option, while two companies relied on it for prospective data collection. The data collected in this manner are excluded from Fig. 1.

During presentation of custom reports for each company participating in the workshop, representatives were most interested in comparisons of their custom retrospective protocol data to the aggregate protocol data of other companies in the dataset. Based on discussions with participating companies and observations during the reporting stage, the research team concluded that the reports were simple to interpret, and few points were unclear across companies. The easy-to-understand nature of the reports allowed for thought-provoking conversations on the measurement and assessment of burden without a need for detailed explanations of

presented tables or figures. This simple format also allowed some participating companies to present reports internally without the need for further explanation or assistance from the research team. This produced helpful feedback from a larger audience and provided additional insights for future iterations of the analysis. Further feedback from companies was collected and is discussed below.

## Discussion

### Analysis of Company Feedback

The participant burden assessment workshop offered attendees an opportunity to gain hands-on experience with the assessment process and to think critically about many aspects of participant burden. The workshop also provided Tufts CSDD and ZS with invaluable feedback regarding both strengths and weaknesses in the data collection process and the algorithm itself. While the participant burden algorithm had already been developed and validated against several measures of protocol performance, this workshop represented a first attempt to apply the algorithm to protocols in a way that would benefit sponsors and provide a wealth of insights.

Workshop attendees were provided a report which contained participant burden score averages overall and by procedure type, and in the case of the prospective protocols, a figure depicting where unique procedures fell on the ‘high burden, high frequency’—‘low burden, low frequency’ spectrum. The use of a score allowed workshop attendees to quickly and objectively understand where their protocols fell in comparison to the benchmark sample, without the potential stigma that may result from the use of a grading system (e.g., A, B, C, etc.). The use of a score also allows for objective feedback when changes are made to a protocol and is granular enough to observe small improvements that may be missed when using broader categorizations. While workshop attendees occasionally debated the inclusion of certain elements in data collection, they appeared to quickly and willingly accept the use of a participant burden score for assessment purposes.

Potentially the greatest benefit of the workshop was a result of performing the exercise itself, rather than the algorithm or resulting scores. While participating in the workshop, attendees were required to think critically and engage in a dialogue regarding participant burden. Workshop attendees had to consider burden that resulted from not only the procedures performed, but also the impact on participants’ daily lives. Individual reports to companies were typically filled with conversations about aspects of protocols that affect perceived burden. Often attendees would discuss whether they thought certain protocol characteristics should be counted in the assessment, or alternative ways to measure the burden resulting from the characteristic. Other conversations centered around the trade-offs of including certain procedures in a protocol—the inclusion of a procedure may add little in the way of data relevant to the current protocol but could provide data that would be useful in future studies, raising the debate of whether that additional data warranted the increase in burden. While the meetings contained both praise and criticism for the methodology, attendees frequently engaged in critical thinking and conversation about the burden experienced by trial participants, and how the design of the protocols impacted it. Further, multiple companies have reported that the assessment was later shown to additional teams within their companies and continued to drive conversations on the topic.

After being presented with their custom reports, workshop attendees were given an opportunity to make requests and provide feedback on the process. Two of the most common requests centered around the need for additional context. Company scores were compared to the sample as a whole and not broken out by phase or therapeutic area. Most workshop attendees indicated that the scores would be more useful to them if they were compared to similar protocols—protocols in the same phase, and if possible, indication. Given the smaller sample size for this exercise,

these exact apples-to-apples comparisons were not feasible. Even making comparisons to the same TA resulted in comparison samples that were too small to be reliable. While this criticism is fair, it is also one that can be easily remedied by increasing the database of evaluated protocols. As the number of protocols increases, the sample size of comparators will naturally increase, allowing for more direct comparisons to be made. This is one of the first areas that future work in this area will focus on—increasing the number of evaluated protocols for comparison.

Several workshop attendees also mentioned that a comparison to the standard of care for the target indication would be useful. As with the therapeutic area and indication breakouts mentioned above, this would allow for an apples-to-apples comparison of what participants would experience within the trial to what they would experience if they were not enrolled. It would also provide context for some particularly high- or low-burden protocols. Indications with high-burden standards of care would be expected to have trials that also have high protocol burdens. In the case of this workshop, one of the most burdensome protocols had a single procedure performed hundreds of times, however, the workshop attendee that provided the protocol indicated that particular procedure was only performed about 25% more often than a patient with the given indication would experience on a daily basis. While the total burden caused by this procedure was very high, when taking the standard of care for the indication into consideration, the additional burden may be comparably low.

The comparison to a standard of care is harder to solve than simply increasing the size of the established database. In essence, it would require establishing a second database of standards of care for all (or at least the most common) indications for which treatments are being developed. The burden scores could then be used as needed to compare to new protocols. Alternatively, a worksheet similar to the protocol worksheet could be developed for standards of care, and when a sponsor company is interested in a protocol evaluation, they could complete both forms, providing the needed comparator. It could be argued, however, that neither of these are necessary, as the source of burden, whether it is from the protocol or the disease itself, is a somewhat moot point. Participants will be experiencing that burden regardless, and the goal of these evaluations should be to reduce the burden experienced by participants as much as possible, not simply to minimize the difference between participating in the trial and not participating.

Although a burden score for standards of care may not be beneficial when trying to reduce the burden of trial participation, it may help inform another area of trial performance—dropout rate. The previous stages of the participant burden algorithm development were unable to identify a significant association between the burden score and participant

dropout rates, and again in this latest round, correlational analysis did not reveal a significant correlation between the two. This lack of correlation could in part be due to the sample size, however throughout the workshop, several attendees raised the topic of dropout rates and offered various possible ‘missing links’ between participant burden and dropout rate. One suggestion was that data on standard of care by indication may shed some light on the matter if dropout rate is at least partially driven by the difference in burden between trial participation and the standard of care. It follows that a participant burden algorithm that does not account for standard of care may not identify existing associations between higher dropout rate and higher participant burden if standard of care procedures are comparable to those of the trial. Therefore, accounting for burden beyond standard of care for individual indications may reveal underlying causes of dropout. Identifying key differences between the standard of care and the experimental treatment could also isolate potential benefits and drawbacks of the new course of treatment.

Workshop attendees also raised the possibility that dropout rate may be related to the perceived benefit of the new treatment. Participants enrolling in a trial may be significantly more motivated to continue with the trial if the treatment represents a substantial improvement over existing treatments, or if there is no treatment available at all. Participants enrolled in trials for new treatments that could extend life may be willing to undergo more burden than those enrolled in a trial for a new treatment intended to alleviate a mild symptom. Similarly, workshop attendees suggested that the side effects of new and existing treatments may play a role in dropout rates. Participants may discontinue not because of the burden experienced but because the new treatment causes more side effects than the old, or they may be willing to undergo more burden because the new treatment has fewer side effects than existing treatments. These potential factors are not captured within a burden score and will need to be further explored in future research.

Another area for future refinements to the participant burden assessment process includes increasing the granularity of the 58 common procedures. Much of the list of procedures functioned well for the recording and estimating of burden, however some procedure types, particularly the questionnaires, may require revision. Workshop attendees pointed out that while the worksheet accounts for differences between in-clinic questionnaires, phone interviews, electronic questionnaires, and others, the length of these questionnaires can range widely. Assigning the same burden to a 5-question questionnaire and a 40-question questionnaire simply because they are both conducted at the clinic will inherently lead to some inaccuracy in the estimated burden. Future refinements to the assessment process should look to (a) evaluate the burden of the most common questionnaires individually, or (b) estimate a per question, or potentially a

time-based (per 5 min, per 10 min, etc...), burden score that can be applied to any questionnaire.

## Barriers to Data Collection

A wealth of variables may contribute to participant burden that are difficult to measure based on trial data typically collected by sponsors. For example, distance to clinic for the participant—in other terms, how far the participant must travel for treatment visits—was identified in the previous stage of algorithm development as a contributor to burden. According to participating sponsors, however, sites do not typically collect this information, and even when it is collected, confidentiality issues prevent this data from being shared with sponsors. When sponsors were unable to provide an estimate of distance traveled, the research team utilized a proxy of participant density (number of participants enrolled/number of sites randomizing participants) to estimate a range for this variable, with higher participant density indicating increased distance to the clinic. This is an inexact estimate, however, and further research and data collection efforts are needed to accurately estimate the effect of travel distance on participant burden.

The need to arrange childcare is another variable that is difficult to approximate based on available data. This often depends on many different personal variables unique to the participant, which may be tied to the age of the participant and their children, income level, or the availability of a secondary caregiver (parents, partner, or other support system). The need to arrange childcare has been estimated by the research team as inflating the overall burden by an additional 5%; but if the presence of participants with this need within the trial is unknown, this inflation percentage is difficult to apply accurately to measure participant burden. Other variables, such as the need to miss work or school, have similar personal factors that make measuring the effect on overall burden difficult.

## Recommendations for Future Studies

When considering ways to decrease the participation burden of a protocol, one factor sponsors may need to consider, beyond the frequency and burden of individual unique procedures, is the necessity and relevance of the data collected by the procedure. Some procedures are essential to the trial and support primary endpoints. These procedures cannot be removed from the protocol, and it is likely the number of times they are performed cannot be reduced. Other procedures may support non-essential endpoints (i.e., endpoints that are exploratory for future studies). For these procedures, it is important for protocol authors to weigh the benefits of this additional data against the increased burden placed on the participants. In some cases, it may be worth reducing



the number of times the procedure is performed, or even removing it from the protocol entirely. Future versions of the data collection worksheet may include classifications of procedures so that the relevance of the procedures, in addition to their frequency and burden score, can be considered when attempting to reduce the participation burden.

Although the participant burden algorithm has been validated against several measures of protocol performance [2], future research could look to validate the score against measures of perceived burden taken directly from trial participants as well. The use of a short, post-participation survey with simple ratings, i.e., a 10-point scale, could be used to show that participants perceive the protocols with lower scores as less burdensome. Further, this could be used to estimate the decrease in burden score necessary for participants to perceive a difference in burden.

Several limitations of the algorithm have already been discussed, but the study itself also faced limitations. Primarily, the protocol sample was not random. Workshop attendees were asked to provide data on several protocols which were included in previous, unrelated research. As such, some amount of bias may have been introduced to the sample. Further, the protocols analyzed were provided by only 8 large sponsor companies and so the protocols may not be representative of the field, or of protocols written by small or mid-sized companies.

A final limitation is a result of the changing field. This workshop included protocols from 2019 and before. Since the onset of the global COVID-19 pandemic, the field of clinical trials has seen some dramatic changes, especially regarding the use of decentralized trials. Virtual trial methodology was a variable within the worksheet but was rarely used by the participating sponsors. Were a similar workshop to be conducted on trials from 2020 and 2021, this design would likely see an increase in use, meaning that the protocols in our sample may not be representative of the current landscape. Regardless, they can still be viewed as an established, pre-pandemic baseline, and confirm the applicability of our methodology.

Companies that aim to utilize a participant burden algorithm should ensure that data collection is performed by dedicated staff with time allotted specifically for this purpose. If dedicated staff are not senior employees, a superior should be assigned as a direct resource to ensure accuracy of data. This will encourage increased efficiency in data collection, and ultimately improve the results of the burden assessment. Companies are likely to see the greatest benefit from a participant burden assessment if it is conducted during protocol planning. During this time feedback from the assessment can be incorporated and acted upon, resulting in a lower burden of participation.

Consistent communication with the team performing the assessment is also an important element of this process, to

clarify any confusion on the part of the company representatives and prevent inaccuracies in the data, which are not always distinguishable by the assessment team following submission of the data collection worksheet. Additionally, provision of the full protocol and amendments along with the completed data collection worksheet allows the assessment team to identify inaccuracies not clarified prior to the submission of data. In the future, efforts to standardize and digitize protocols may reduce the effort needed to collect data and improve the accuracy of the data collected. During the review of the participant burden assessment results, it is important to include team members that have enough knowledge of the protocol to discuss the feasibility of recommendations made (i.e., whether it is possible to reduce the number of times a procedure is performed) as well as team members in a position to ensure that the changes are implemented.

Long term, sponsors should consider strategies to collect certain variables pertaining to participant burden which are not currently gathered by sites (e.g., average visit time, participant pickup or travel reimbursement offered, previous clinical trial experience, distance from center, need to arrange childcare, and whether the clinical trial protocol required the participant to miss work or school). Sites could even contribute to recording of standard of care treatment data. When collection of these variables is possible, the precision of the algorithm and the benefit of comparisons to prospective protocols will increase.

As we continue to refine our participant burden algorithm instrument, the expansion of our database to allow apples-to-apples comparisons between protocols of the same therapeutic area and indication will be a priority. Additionally, as several non-procedural variables are not currently collected by sponsors, and the development of strategies to collect these variables may take time, validating and revising existing proxies for these variables is essential to the optimization of the algorithm. Further, continued exploration of the relations between certain protocol characteristics and various measures of protocol performance is necessary.

Participant burden is an intricate subject with many contributing factors. Though some of these factors may be too unique to each participant to accurately account for across protocols, many are measurable variables that can be incorporated into a participant burden algorithm. This first attempt to apply our participant burden algorithm to protocol data provided by sponsors has shown the value of this tool in measuring burden and assessing areas for improvement in prospective studies. However, sponsor feedback collected throughout this project indicated that there is still a considerable amount of refinement needed to enhance accuracy and expand applicability of the methodology. Once these refinements have been made, sponsors will be able to apply the participant burden algorithm to evaluate and mitigate



burden in upcoming studies with more accuracy, and on a larger scale.

### Acknowledgements

The authors of this paper wish to thank ZS for the funding provided for this research. They would also like to thank Diana Wang for her contribution to this manuscript.

### Author Contributions

ZS: Substantial contributions to the conception and design of the work and the acquisition, analysis, or interpretation of data for the work. Drafting the work or revising it critically for important intellectual content. Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. EB: Substantial contributions to the conception and design of the work and the acquisition, analysis, or interpretation of data for the work. Drafting the work or revising it critically for important intellectual content. Final approval of the version to be published. CC: Substantial contributions to the conception or design of the work and drafting the work or revising it critically for important intellectual content. AB: Substantial contributions to the conception or design of the work and drafting the work or revising it critically for important intellectual content. BQ: Substantial contributions to the conception or design of the work. KG: Substantial contributions to the conception and design of the work. Drafting the work or revising it critically for important intellectual content. Final approval of the version to be published.

### Funding

This study was supported, in part, by a research Grant from ZS.

### Declarations

#### Conflict of interest

Authors have no conflicts to report.

### References

1. Getz K, Sethuraman V, Rine J, et al. Assessing patient participation burden based on protocol design characteristics. *Ther Innov Regul Sci.* 2019;54:598–604. <https://doi.org/10.1177/2168479019867284>.
2. Smith Z, Wilkinson M, Carney C, et al. Enhancing the measure of participation burden in protocol design to incorporate logistics, lifestyle, and demographic characteristics. *Ther Innov Regul Sci.* 2021;55(6):1239–49. <https://doi.org/10.1007/s43441-021-00336-2>.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.