DIA

ORIGINAL RESEARCH

Impact of Using A Mixed Data Collection Modality on Statistical Inferences in Decentralized Clinical Trials

Alexandra Curtis[1] · Yongming Qu[1]

Abstract

Background Decentralized clinical trials offer the promise of reduced patient burden, faster and more diverse recruitment, and have received regulatory support during the COVID-19 pandemic. However, lack of data accuracy or data validation poses a challenge for fully decentralized trials. A mixed data collection modality where onsite measurements are collected at key time points and decentralized measurements are taken at intermediate time points is attractive operationally. To date, the impact of decentralized measurements (which could presumably be less accurate) taken at intermediate time points on statistical inference on the primary or other key time points has not been evaluated.

Methods In this article we evaluate the estimation and statistical inference for three scenarios: (1) all onsite measurements, (2) a mixture of onsite and decentralized measurements, and (3) all decentralized measurements, in the setting of a chronic weight management trial. We consider scenarios where decentralized measurements have additional within- and between-subject variabilities and/or bias.

Results In the mixed modality setting, simulation studies showed that the estimation and inference for the key time points with onsite measurements have good properties and are not impacted by the additional variability and bias from intermediate decentralized measurements. However, estimates for intermediate decentralized time points for the mixed modality and estimates for the all decentralized modality measurements have increased variability and bias.

Conclusion Mixed modality trials can help achieve the benefits of decentralized clinical trials by reducing the number of onsite visits with little impact on statistical inferences for various estimands, compared to traditional (all onsite) clinical trials.

Keywords Estimands · Imputation · Inference · Mixed modality · Remote

Introduction

Decentralized clinical trials (DCTs) are defined by the Clinical Trial Transformation Initiative as "trials executed through telemedicine, mobile/local healthcare providers (HCPs), and/or mobile technologies." [1] A DCT does not need to be completely remote [1, 2], and others have proposed metrics for the degree of decentralization for a clinical trial which are based on how much a study depends on specific sites and the methods of data collection used [2]. In this article, we use the term decentralized to denote data not collected through the traditional approach (i.e., during in-person clinical visits). Decentralized measurements may include measurements collected at home by patients or health workers, through a remote electronic device (e.g., electronic clinical outcome assessment), a mobile device, or at a local medical facility.

Prior to the COVID-19 pandemic, there was interest in DCTs due to the potential reduction in time and cost associated with patient travel which may improve patient retention [1, 3]. Additional potential benefits include faster and more diverse patient recruitment due to fewer geographical restrictions, especially for rare diseases [1, 3, 4]. Technological advances including the wide use of mobile devices and increasing popularity in the adoption of wearable devices also facilitate DCTs [1, 5, 6].

During the COVID-19 pandemic, interest in decentralization turned into necessity and use of decentralization in clinical trials accelerated quickly due to patient safety concerns about traveling to clinical trial sites or temporary closure

✉ Alexandra Curtis
curtis_alexandra@lilly.com

1   Department of Statistics, Data and Analytics, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, USA

of some sites [7]. This public health emergency triggered regulatory support for alternative methods of data collection with an emphasis on maintaining patient safety, which helped temporarily facilitate trial decentralization [3, 7–9].

Despite the advantages and recent regulatory support for DCTs, these trials still present many challenges. For example, there is uncertainty about which of the additional regulatory flexibilities will be allowed once the COVID-19 pandemic ends [3, 7]. Additional legal complications in the USA include restriction of telehealth visits over state lines, which may require having primary investigators with a medical license in each patient's home state. Delivery of study medications directly to participants may also be difficult for practical reasons, due to some state laws, the need for preservation of patient privacy during delivery, and potential need for refrigeration [1, 10]. From a patient safety perspective, there is also concern about how sponsors and Institutional Review Boards will monitor and enforce safety requirements on decentralized sites, as well as the accuracy of data collected to monitor adverse events [8, 9].

Alternative data collection in DCTs may lead to data quality concerns including whether the data is traceable/auditable, how to securely store or transmit patient data, whether the data will have increased variability or biases, and how updates to the software or hardware used to collect the data will impact longitudinal studies [6, 11]. The United States Food and Drug Administration supports some decentralized data collection for trials ongoing during the pandemic [12], but it may take time before the effects of decentralization on data quality are well understood.

In some cases, local, home, or virtual data collection may not be as reliable as data collected onsite. Some virtual data collection tools and methods require validation. Despite the sponsor's best effort to validate the devices and tools for decentralized data collection, the quality of the decentralized measurements may not be as consistent as onsite measurements. Using body weight as an example, even if the sponsor provides a standardized weight scale and detailed instructions, patients generally do not have the necessary training to be fully compliant to the procedure.

One approach to addressing some of these challenges while planning a new DCT is to utilize a hybrid approach for decentralization by conducting a few critical visits on site and all other data collection and visits in a decentralized manner [1]. This approach was recommended for one DCT study by a regulatory agency [Unpublished Communication]. In this article, we use clinical trials in chronic weight management (CWM) as an example to understand the impact of a fully decentralized approach and a hybrid approach on the estimates of mean response and the treatment difference in the primary outcome under various scenarios including increased variability and bias for decentralized measurements. Section 2 describes the clinical trial

setting and the data generating models, Section 3 provides simulation results, and in Sections 4 and 5 we summarize our findings.

Materials and Methods

Simulation Setting

We consider a 52-week trial comparing an experimental treatment with placebo for the purpose of CWM for participants with obesity, defined by baseline body mass index (BMI) ≥ 30 kg/m^2 or BMI ≥ 27 kg/m^2 with at least 1 weight-related comorbidity. The primary analysis variable is the change in body weight from baseline to 52 weeks. We assume body weight is measured at baseline and every 4 weeks after randomization until week 52, and the study enrolls 600 participants randomly assigned to placebo or an experimental treatment with a 1:1 randomization ratio.

The onsite weight measurement for subject j ($1 \leq j \leq n_i$) at time point k ($k = 0, 1, \ldots, 13$) on treatment i ($i = 0$ for the control group and $i = 1$ for the experimental treatment group) is modeled as

$$Y_{ijk} = \mu_{ik} + s_{ij} + e_{ijk}, \tag{1}$$

where μ_{ik} is the population mean response for treatment group i at time point k, $s_{ij} \sim$ NID $(0, \sigma_s^2)$, $e_{ijk} \sim$ NID $(0, \sigma_e^2)$, where s_{ij} and e_{ijk} are independent and NID stands for "normally independently distributed". We assume the decentralized measurement Y_{ijk}^* follows

$$Y_{ijk}^* = (1 + r)(\mu_{ik} + s_{ij} + e_{ijk} + s_{ij}^* + e_{ijk}^*), \tag{2}$$

where $s_{ij}^* \sim$ NID $(0, \sigma_s^{2*})$ is the additional between-subject variability, $e_{ijk}^* \sim$ NID $(0, \sigma_e^{2*})$ is the additional within-subject variability, s_{ij}^*, s_{ij}, e_{ijk}^* and e_{ijk} are all mutually independent, and r ($-1 < r < 1$) is the proportion of bias in the decentralized measurement.

We consider 3 clinical trial settings:

1. All onsite—Body weight is measured on site for all time points.
2. Mixed—Body weight is measured on site for Week 0 (baseline), Week 24, and Week 52, and decentralized measurements are used for all other time points.
3. All decentralized—Decentralized body weight measurements are used for all time points.

For decentralized measurements, we consider 3 scenarios: (A) additional within-subject variability but without additional between-subject variability and bias ($\sigma_e^{2*} > 0$, $\sigma_s^{2*} = 0$, and $r = 0$), (B) additional within- and

between-subject variability but without bias ($\sigma_e^{2*} > 0$, $\sigma_s^{2*} > 0$, and $r = 0$), and (C) additional within- and between-subject variability and additional bias ($\sigma_e^{2*} > 0$, $\sigma_s^{2*} > 0$, and $r \neq 0$).

For the onsite measurements, we use a between-patient variance (σ_s^2) of 300 and a within-patient variance (σ_e^2) of 70. These values are selected based on historical studies. For decentralized measurements, the parameters for additional variability and bias are $\sigma_s^{2*} = 130$, $\sigma_e^{2*} = 30$, and $r = -0.05$. The σ_s^{2*} and σ_e^{2*} values assume an approximate 20% increase in the within- and between-patient standard errors beyond the σ_e^2 and σ_s^2 values from historical studies. Literature reviews have suggested self-reported weight measurements are typically 0.5 to 1 kg lower compared to weight measurements taken on site [13, 14]. We assume an -5% (approximately -5 kg) bias to evaluate the performance of estimators under an extreme situation.

In CWM studies, often the primary endpoint is the percent change in body weight from baseline. Under the data generation models (1) and (2), the theoretic mean percent weight loss is difficult to calculate, so we first use the change in body weight (kg) from baseline as the primary analysis variable and perform analyses to estimate the estimands for the change in body weight. The true mean body weight, the mean change in body weight, and the mean percent change in body weight by treatment group are provided in Table 1. The mean percent change was calculated based on 1,000,000 Monte Carlo simulation samples.

To introduce missing data under the missing at random assumption [15], we first calculate the change in body weight from baseline for each post-baseline visit. Then, we generate dropouts assuming the probability of patients staying in the study at the k-th time point ($k = 2, 3, \ldots, 13$) is:

$$\pi_{ijk} = \frac{1}{1 + \exp\left\{\alpha_0 + \alpha_1\left[\{Z_{ij,k-1}Y_{ij,k-1} + (1 - Z_{ij,k-1})Y_{ij,k-1}^*\} - \{Z_{ij0}Y_{ij0} + (1 - Z_{ij0})Y_{ij0}^*\}\right]\right\}},$$

(3)

where Z_{ijk} is a binary variable indicating whether the measurement at time point k is collected on site. The probability of missingness at each time point depends on the observed intermediate change in weight measurement at the previous time point. This missingness mechanism is ignorable missingness, also called missing at random (MAR) [15]. We choose $\alpha_1 = 1$ to ensure the data is missing at random and α_0 ranging from -17.5 to -24.5 to obtain an overall proportion of patients with missing values at Week 52 of approximately 20%.

Inferences

Under the assumption of missing at random, we estimate the mean change in body weight using a mixed model for repeated measures (MMRM) including factors of treatment, time point, and treatment by time point interaction, with baseline body weight as a covariate. An unstructured variance–covariance structure is used to model the within-subject correlation. Commonly used software packages (SAS and R) as well as theory for (mixed) linear models provide inference conditional on the baseline covariate. Let $\hat{\mu}_{ik}$ denote the estimator for μ_{ik} and $\text{Var}\{\hat{\mu}_{ik}|\bar{X}_{..}\}$ be the variance of $\hat{\mu}_{ik}$ conditional on baseline covariate, where $\bar{X}_{..}$ represents the overall mean for X across treatment groups. By the law of total variance, we have

Table 1 True Mean Body Weight by Treatment Arm

Week	Mean (kg)		Mean CFB (kg)		Mean %CFB	
	Placebo	Treatment	Placebo	Treatment	Placebo	Treatment
0	101	101	0	0	0	0
4	100.5	99.0	-0.5	-2.0	0.3	-1.3
8	99.5	97.5	-1.5	-3.5	-0.8	-2.8
12	99.0	96.0	-2.0	-5.0	-1.3	-4.4
16	98.5	95.0	-2.5	-6.0	-1.8	-5.4
20	98.0	94.0	-3.0	-7.0	-2.3	-6.4
24	98.0	92.0	-3.0	-9.0	-2.3	-8.5
28	98.0	91.0	-3.0	-10.0	-2.3	-9.5
32	98.0	90.5	-3.0	-10.5	-2.3	-10.0
36	98.0	90.0	-3.0	-11.0	-2.3	-10.6
40	98.0	88.5	-3.0	-12.5	-2.3	-12.1
44	98.2	88.0	-2.8	-13.0	-2.1	-12.6
48	98.3	87.0	-2.7	-14.0	-2.0	-13.6
52	98.5	86.0	-2.5	-15.0	-1.8	-14.7

CFB change from baseline; *kg* kilogram.

$$\text{Var}\{\hat{\mu}_{ik}\} = E[\text{Var}\{\hat{\mu}_{ik}|\bar{X}_{..}\}] + \text{Var}[E\{\hat{\mu}_{ik})|\bar{X}_{..}\}].$$

Then,

$$\widehat{\text{Var}}\{\hat{\mu}_{ik}\} = \widehat{\text{Var}}\{\hat{\mu}_{ik}|\bar{X}_{..}\} + \widehat{\text{Var}}[E\{\hat{\mu}_{ik}|\bar{X}_{..}\}]. \tag{4}$$

Linear model packages give the variance estimate $\widehat{\text{Var}}\{\hat{\mu}_{ik}|\bar{X}_{..}\}$. We need to add the second part to obtain the unconditional variance estimation. The second part can be estimated as

$$\widehat{\text{Var}}[E(\hat{\mu}_{ik}|\bar{X}_{..})] = (\hat{\beta}_x)^2\text{Var}(\bar{X}_{..}) = (\hat{\beta}_x)^2 S_x^2/n,$$

where S_x^2 is the sample variance for the baseline observations (across treatment groups) and $n = n_1 + n_2$ is the overall sample size. Since the estimator $\hat{\mu}_{1k} - \hat{\mu}_{0k}$ for the treatment effect does not contain $\bar{X}_{..}$, the conditional and unconditional variance are the same and the direct variance estimation from software packages can be directly used for inference. In the simulation results in Section 3, the inference for the mean in each treatment group is based on the variance estimation in Eq. (4).

Results

We conducted simulation studies based on 5,000 simulated samples for 3 clinical settings (all onsite measurements, measurements with mixed modality, and all decentralized measurements) for scenarios A, B, and C for increased variability and bias. In all scenarios, the proportion of missing values at the 52-week visit is between 13.8 and 15.6% for the experimental treatment arm and 24.3 and 26.6% for the placebo arm.

Figure 1 summarizes the simulation results for Scenario A in which we assume decentralized measurements only increase the residual error [$\sigma_e^{2*} = 30$, $\sigma_s^{2*} = 0$, and $r = 0$ in Model (2)]. When measurements are all collected on site, the bias is minimal and the estimated 95% confidence intervals have a coverage probability of approximately 0.95. For the mixed onsite/decentralized measurement modality, the standard errors at visits with decentralized measurements are larger than the case of all onsite measurements, but at the visits where weights are measured onsite (Weeks 24 and 52), the point estimates and standard errors are similar between the mixed modality and all onsite measurements. For the case of all decentralized measurements, all estimates also have minimal bias but the standard errors are consistently larger than the all onsite measurement scenario.

Figure 2 summarizes the simulation results for Scenario B in which we assume decentralized measurements increase
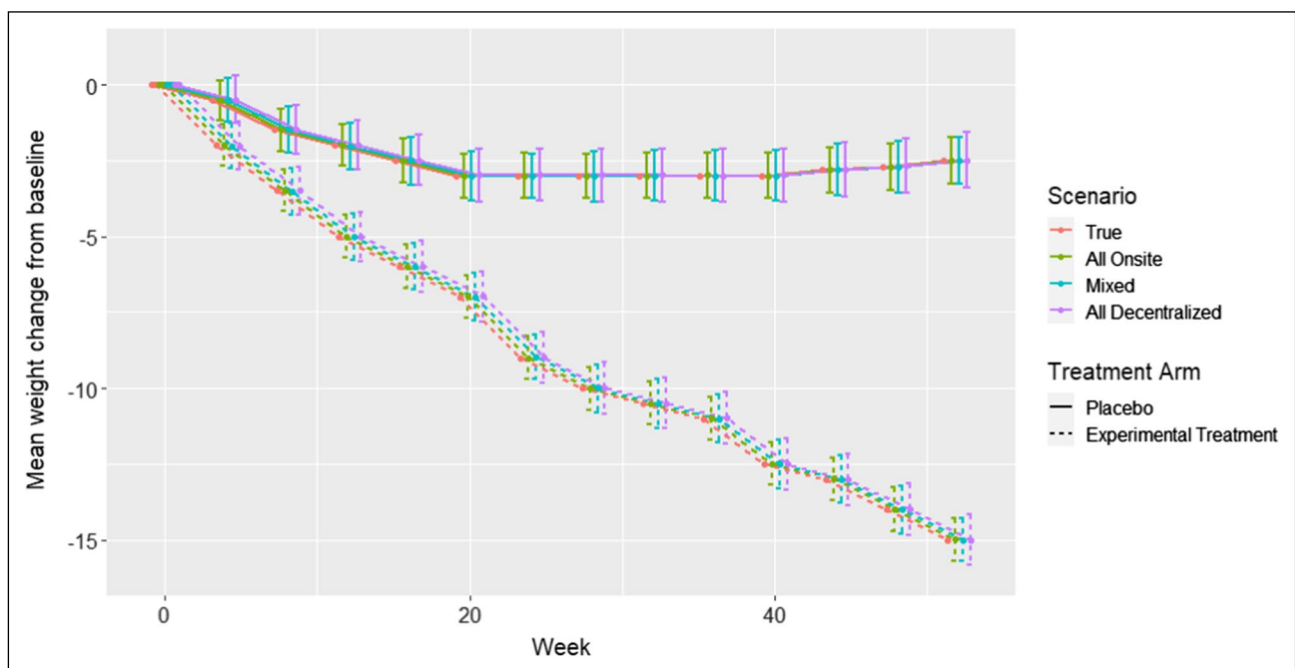

Figure 1 Summary of Simulation Results for the Treatment Difference in the Mean Change in Body Weight from Baseline for the Scenarios with All Onsite Measurements, with a Mixture of Onsite and Decentralized Measurements, and with All Decentralized Measurements Assuming Increased Within-Subject Variability for Decentralized Measurements (Based on 5000 Simulated Samples). The True values indicate the true weight change from baseline for each treatment arm and time point.
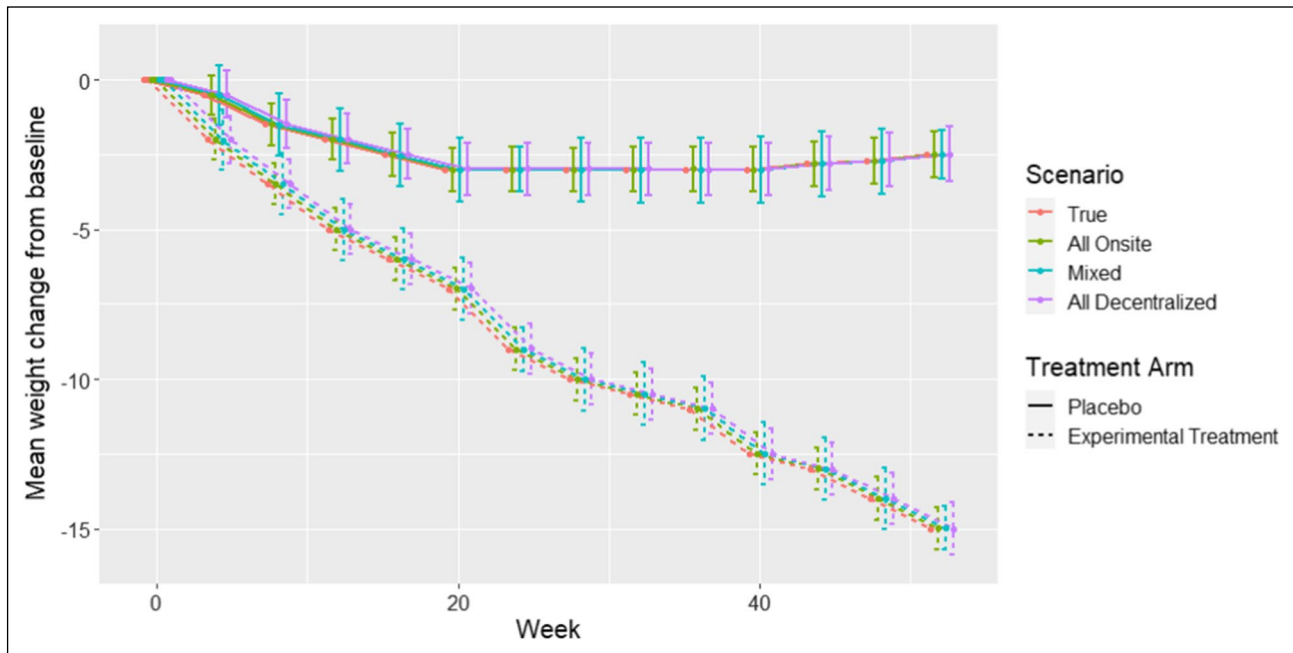
Figure 2 Summary of Simulation Results for the Treatment Difference in the Mean Change in Body Weight from Baseline for the Scenarios with All Onsite Measurements, with a Mixture of Onsite and Decentralized Measurements, and with All Decentralized Measurements Assuming Increased Within- and Between-Subject Variability for Decentralized Measurements (Based on 5000 Simulated Samples). The True values indicate the true weight change from baseline for each treatment arm and time point.

the residual error variability and between-subject variability [$\sigma_e^{2*} = 30$, $\sigma_s^{2*} = 130$, and $r = 0$ in Model (2)]. The results are generally similar to those for Scenario A (Fig. 1), with slight increases in standard errors for the decentralized measurements in the mixed modality.

Figure 3 summarizes the simulation results for Scenario C in which we assume the decentralized measurements not only increase the within- and between-subject variabilities but also introduce a bias [$\sigma_e^{2*} = 30$, $\sigma_s^{2*} = 130$, and $r = -0.05$ in Model (2)]. For mixed modality measurements, the biases and standard errors in the estimated mean change in body weight from baseline to 24 and 52 weeks are similar to the case of only onsite measurements, but the biases are larger for other time points. Recall that in the mixed modality scenario the measurements at baseline, Week 24 and Week 52 were collected on site, so we expect to see no bias in the change from baseline estimates to 24 and 52 weeks, but the bias for the change in body weight for intermediate time points is amplified due to the mismatch of data collection methods at baseline and the intermediate time points. For the case of all decentralized measurements, there is a clear bias in all time points but the bias is smaller than the decentralized measurement time points in the mixed modality scenario. This is likely because the bias is largely canceled out when calculating the change from baseline when both baseline and post-baseline measurements are decentralized. Note the bias

cancelation in the all decentralized modality is specific to our simulation scenarios. If the effect of bias cancelation is smaller, the bias for the all decentralized scenario will be larger.

Table 2 summarizes the bias, standard deviation, the mean of the standard error over the simulation iterations, and the coverage probability for the 95% confidence interval for the treatment difference in the mean change in body weight from baseline to 12, 24, 36, and 52 weeks. Data for other intermediate time points are not shown here, as the results were similar to those at 12 and 36 weeks. Note for the mixed modality setting, onsite measurements are collected at baseline, 24, and 52 weeks only. Similar to the results for each treatment group, for the mixed data collection modality the estimates for the treatment difference are very similar to those for the non-DCT (traditional) setting for 24 and 52 weeks, have larger variability at 12 and 36 weeks when decentralized measurements have increased variabilities, and are biased at 12 and 36 weeks when decentralized measurements have measurement bias. Note the bias is much smaller for the treatment difference compared to individual treatment group results. This is likely because some bias is canceled out during calculation of the treatment difference. For the case of all decentralized measurements, estimates for the treatment difference have increased variability at all time points and are slightly biased when the decentralized measurements are biased.
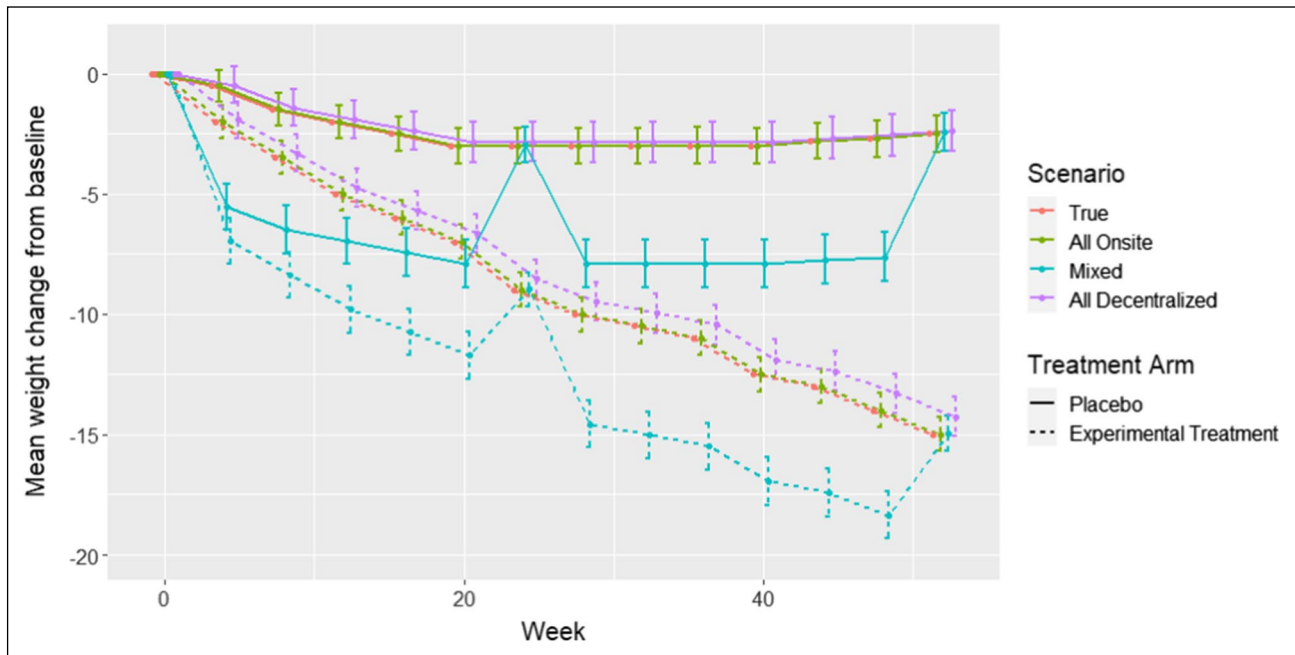
Figure 3 Summary of Simulation Results for the Treatment Difference in the Mean Change in Body Weight from Baseline for the Scenarios with all Onsite Measurements, with a Mixture of Onsite and Decentralized Measurements, and with All Decentralized Measurements Assuming Increased Within- and Between-Subject Variability, and Bias for Decentralized Measurements (Based on 5000 Simulated Samples). The True values indicate the true weight change from baseline for each treatment arm and time point.

As mentioned previously, primary outcomes from CWM trials are commonly presented as percent change in weight from baseline. Therefore, we carried out analogous simulation studies based on percent change from baseline and the results can be found in the Supplemental Information online. Overall, the coverage probabilities for the percent change results are slightly lower for all cases. This is likely because the percent change data does not satisfy all the assumptions of the MMRM model used (i.e., the percent change values have heterogeneous variance and non-normality). However, the trends in increasing standard deviations and bias at decentralized visits, as well as lower coverage probabilities in the scenario where bias is present at decentralized visits, are all consistent with the results discussed previously when the absolute change in weight from baseline is the primary outcome of interest.

We did not simulate the case under the null hypothesis of no treatment difference. Based on the theory of linear models, the coverage probability does not change when the true mean treatment difference changes for linear models. Therefore, the coverage probability in Table 2 can be used gauge the type I error rate (type I error is equal to 1 minus the coverage probability). Referring to this table for the Scenario C results (which correspond to Fig. 3), the type I error rate at 52 weeks in the mixed modality is approximately 5% and the type I error rate in the all decentralized modality is 9%. Due to the inflated type I error and biased estimates

in the all decentralized modality at critical time points, the mixed modality is a reasonable approach to increase the number of clinical trial visits which are decentralized while maintaining the validity of the inference at key time points if the decentralized data measurements are not fully validated.

Discussion

DCTs are drawing more interest due to the convenience for patients and easier access for traditionally under-represented patients. One challenge for DCTs is the validity of decentralized measurements (collected at home, virtually, or at local facilities). For measurements collected by patients, even if standardized devices and detailed instructions are provided to patients, they may not be able to closely follow the instructions as they are not trained technicians or nurses. In addition, completely virtual clinical trials can be challenging as patients and physicians may prefer to have some face-to-face interactions with sites and certain measurements can be more reliably obtained on site versus at home or locally. A mixed modality of on site and decentralized visits may be attractive because they can take advantage of the benefits of DCTs (less travel time and more access to under-represented patient populations) while maintaining the benefit of face-to-face interactions between healthcare providers and patients, and accurate measurements taken at traditional site-based

Table 2 Summary of simulation results for the difference between placebo and experimental treatment arm in the difference in weight from baseline—comparing the all onsite scenario to scenarios with decentralized measurements at select time points (based on 5000 simulated samples)

Scenario	Modality	Week[a]	True	DCT Scenario Results			
				Bias	SD	SE	CP
All onsite		12	− 3.0	− 0.002	0.957	0.954	0.951
		24	− 6.0	− 0.004	0.990	0.974	0.942
		36	− 8.0	− 0.007	0.993	0.986	0.950
		52	− 12.5	0.015	1.002	0.996	0.947
Within patient variability (Scenario A)	Mixed	12	− 3.0	0.010	1.060	1.060	0.951
		24	− 6.0	0.015	0.997	0.975	0.944
		36	− 8.0	0.003	1.115	1.101	0.946
		52	− 12.5	0.002	1.004	0.999	0.950
	All Decentralized	12	− 3.0	− 0.022	1.113	1.120	0.952
		24	− 6.0	− 0.006	1.148	1.145	0.950
		36	− 8.0	0.012	1.155	1.160	0.950
		52	− 12.5	− 0.013	1.180	1.173	0.947
Within and between patient variability (Scenario B)	Mixed	12	− 3.0	0.011	1.437	1.423	0.946
		24	− 6.0	0.001	0.987	0.990	0.949
		36	− 8.0	0.037	1.462	1.450	0.942
		52	− 12.5	0.021	1.025	1.012	0.946
	All Decentralized	12	− 3.0	− 0.018	1.131	1.138	0.953
		24	− 6.0	− 0.003	1.168	1.163	0.949
		36	− 8.0	0.015	1.175	1.178	0.949
		52	− 12.5	− 0.010	1.198	1.191	0.946
Within and between patient variability and bias (Scenario C)	Mixed	12	− 3.0	0.137	1.345	1.351	0.953
		24	− 6.0	0.005	0.994	0.989	0.950
		36	− 8.0	0.428	1.363	1.378	0.942
		52	− 12.5	− 0.030	1.016	1.013	0.946
	All Decentralized	12	− 3.0	0.148	1.077	1.081	0.953
		24	− 6.0	0.306	1.106	1.105	0.940
		36	− 8.0	0.409	1.125	1.118	0.933
		52	− 12.5	0.627	1.145	1.131	0.912

CP coverage probability of the 95% confidence interval, *SD* standard deviation of estimates of the mean, *SE* mean standard error estimates of the mean, *True* the true mean value.

[a] For the onsite/decentralized mixed measurements models, weeks 24 and 52 were site visits.

clinical trials. In Eli Lilly and Company, most new diabetes clinical trials utilizing DCT capabilities use a mixed modality, in which onsite measurements are collected at key time points (e.g., baseline and the primary time points).

The impact of the mixed DCT modality on statistical inference (bias, variance, and coverage probability of the confidence interval) has not been studied previously. In this article, we evaluated the impact of the mixed DCT modality on statistical inference through simulation for longitudinal continuous outcomes. Since we assume the measurements collected at baseline and primary time points are taken onsite, the DCT measurements at intermediate time points would not impact the primary analysis if there were no missing values. In our simulation, we evaluate the performance of the MMRM under the assumption of MAR and our results show the modality at intermediate outcomes has little impact on the estimation of treatment effect at key time points.

The MMRM analysis is often used to estimate treatment effects when the hypothetical strategy is used to handle intercurrent events and missing or unobserved values are considered missing at random. Another important strategy commonly used in handling intercurrent events is the treatment policy strategy. To estimate the treatment effect under the treatment policy strategy, imputation methods (e.g., retrieved dropout [16] and reference-based imputation [17] methods) based on the pattern mixture model may be used. Since retrieved dropout and jump-to-reference imputation methods only use the baseline and primary time point (for the primary analysis), the modality of the intermediate measurements will not impact these 2 imputation methods. For copy-reference and return-to-baseline [18] imputation methods, intermediate outcomes are needed to estimate the within-subject variance–covariance structure for the reference arm and/or treatment arm, which requires
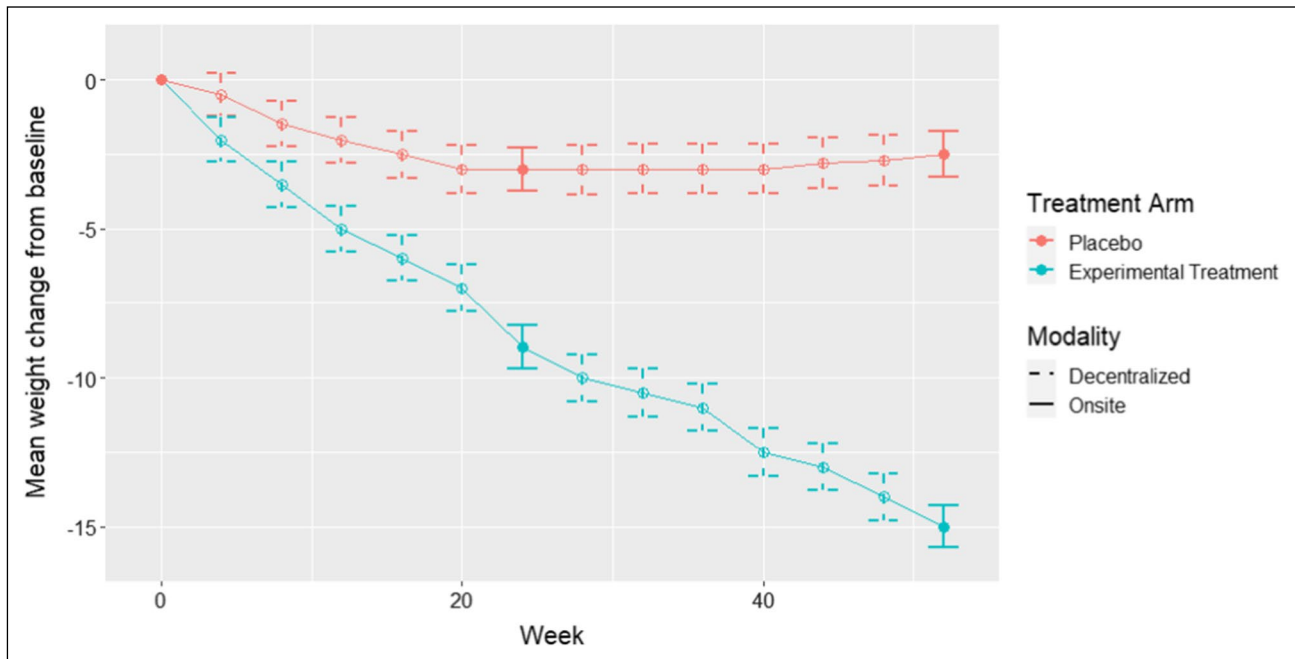
Figure 4 Visualization of the Mean Change in Body Weight Over Time When Mixed-Modality Data Collection Methods are Used in a Decentralized Clinical Trial

the assumption of ignorable missingness for hypothetical outcomes. Based on simulation results in the "Results" section under the missing at random assumption, it is reasonable to conclude the mixed modality also has little impact on estimation using copy-reference and return-to-baseline imputation methods. However, we note the different modality for data collection at intermediate time points may impact the amount of intercurrent events or missing values which may affect estimation for the estimands associated with the treatment policy strategy and aforementioned imputation methods, but we expect such an impact on the final estimation results to be small. Overall, data collected in the mixed modality setting can be used to handle missing data in a wide variety of ways and the best approach to handling missing data in a particular study depends on the reason for missing data and the estimand being used.

One drawback of using the mixed modality approach is the potential for high bias at non-critical time points as demonstrated in Fig. 3. An alternative approach which may help address this drawback is to collect both onsite and decentralized measurements at baseline and use the decentralized baseline measurement to calculate the change from baseline at decentralized time points. This is a more complicated option which assumes that the bias observed in a decentralized measurement is consistent across time points. This could also potentially increase the challenge of statistical analysis, e.g., which (onsite or decentralized) baseline value should be included in the analysis model. We do not recommend this approach.

Another challenge in using the mixed DCT modality is understanding the mean response over time since onsite and DCT measurements may not be "comparable". For most measurements, the difference between measurement modalities may not be large. To present the estimates from onsite measurements taken at key time points and estimates from DCT measurements at intermediate time points, one can use different line type and point symbols to differentiate the measurement modalities (see Fig. 4 for an illustration). Additionally, for a variable that is collected through a decentralized approach for the first time, it may not be a bad idea to collect both decentralized and onsite measurements at some time points so that the difference between the 2 data collection modalities can be assessed.

Conclusion

In conclusion, clinical trials with mixed decentralized and onsite measurement modalities are an attractive option to reduce the burden of patient travel to the trial site while maintaining the advantages of traditional clinical trials. This article shows a mixed modality approach has little impact on analysis results at the primary time point(s) for various estimands, even if the decentralized measurement devices or procedures are not fully validated. Therefore, more clinical trials with mixed modality should be conducted to attract traditionally under-represented patient populations, reduce

patients' burden, and reduce the cost of clinical trials while maintaining the quality of inference for key study objectives.

Author Contributions

Both authors have made substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; and drafted the work or revised it critically for important intellectual content; and made the final approval of the version to be published; and agree 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 or resolved.

Funding

Declarations

Conflict of interest

The authors are shareholders of Eli Lilly and Company.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1007/s43441-022-00416-x.

References

1. Apostolaros M, Babaian D, Corneli A, et al. Legal, regulatory, and practical issues to consider when adopting decentralized clinical trials: recommendations from the clinical trials transformation initiative. *Therap Innov Regul Sci*. 2020;54:779–87.

2. Khozin S, Coravos A. Decentralized trials in the age of real-world evidence and inclusivity in clinical investigations. *Clin Pharmacol Therap*. 2019;106:25–7.

3. Kadakia KT, Asaad M, Adlakha E, et al. Virtual clinical trials in oncology—overview, challenges, policy considerations, and future directions. *JCO Clin Cancer Inf*. 2021;5:421–5.

4. Sommer C, Zuccolin D, Arnera V, et al. Building clinical trials around patients: evaluation and comparison of decentralized and conventional site models in patients with low back pain. *Contemp Clin Trials Commun*. 2018;11:120–6.

5. Gerke S, Shachar C, Chai PR, et al. Regulatory, safety, and privacy concerns of home monitoring technologies during COVID-19. *Nat Med*. 2020;26:1176–82.

6. Walton MK, Cappelleri JC, Byrom B, et al. Considerations for development of an evidence dossier to support the use of mobile sensor technology for clinical outcome assessments in clinical trials. Contemporary Clinical Trials. 2020;91.

7. Krofah E, Schneeman K. Lessons learned from COVID-19: are there silver linings for biomedical innovation? 2021. https://milkeninstitute.org/sites/default/files/reports-pdf/MI_SilverLining_012521.pdf.

8. United States Department of Health and Human Services Food & Drug Administration. Conduct of clinical trials of medical products during the COVID-19 public health emergency - guidance for industry, Investigators, and Institutional Review Boards; 2021. https://www.fda.gov/media/136238/download. Accessed 20 Sept 2021.

9. United States Department of Health and Human Services Food & Drug Administration. Statistical considerations for clinical trials during the COVID-19 public health emergency—guidance for industry; 2020. https://www.fda.gov/media/139145/download. Accessed 20 Sept 2021.

10. TransCelerate BIOPHARMA Inc. Beyond COVID-19: modernizing clinical trial conduct; 2020. https://www.transceleratebiopharmainc.com/wp-content/uploads/2020/07/TransCelerate_Beyond-COVID19_Modernizing-Clinical-Trial-Conduct_July-2020.pdf.

11. Carroll G, Sullivan K, Traverso K, et al. *Harnessing safety data from wearable devices*. Hermitage, TN: Deloitte Development LLC; 2018.

12. U S Food & Drug Administration. Advancing oncology decentralized trials; 2021. https://www.fda.gov/about-fda/oncology-center-excellence/advancing-oncology-decentralized-trials. Accessed 20 Sept 2021.

13. Gorber SC, Tremblay M, Gorber B. A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. *Obesity Rev*. 2007;8:307–26.

14. Maukonen M, Mannisto S, Tolonen H. A comparison of measured versus self-reported anthropometrics for assessing obesity in adults: a literature review. *Scand J Public Health*. 2018;46:565–79.

15. Rubin DB. Inference and missing data. *Biometrika*. 1976;63(3):581–92.

16. European Medicines Agency Committee for Medicinal Products for Human Use (CHMP). Guideline on missing data in confirmatory clinical trials. EMA/CPMP/EWP/1776/99 Rev. 1.; 2010.

17. Carpenter JR, Roger JH, Kenward MG. Analysis of longitudinal trials with protocol deviation: a framework for relevant, accessible assumptions, and inference via multiple imputation. *J Biopharm Stat*. 2013;23(6):1352–71.

18. Qu Y, Dai B. Return-to-baseline multiple imputation for missing values in clinical trials. *Pharmac Stat*. 2022;21:641–53.