Manuscript RSM-06-2021-0105 investigates the bias issue relating to the complete-case analysis and the shifting-case analysis for meta-regression in the presence of missing covariates. The main findings highlight that the magnitude of bias critically depends on whether or not the probability of observing a covariate is dependent of effect sizes, which represent outcomes in the setting of meta-regression for the complete-case analysis. This result is consistent with existing literature reporting impact of missing covariates on complete-case analyses. I have the following comments for the authors to consider:

1. Section 2. In the substance abuse interventions example, it is mentioned that "Often the same group (typically the control group) in a study was used in multiple contrasts...". For a control group, how would a covariate indicating high- versus low-intensity intervention defined? It seems that this would be undefined, rather than missing?

2. Section 2. Page 3. The discussion contrasting coefficients for X_1 in complete-case analysis (Column 1 in Table 2) and the coefficient for X_1 in Shifting-case Group 1 (Column 2 in Table 2; and similarly for X_2) is misleading. They are not the same parameter and are usually not expected to the same unless under some special cases. They are not comparable. Whether they are close or different does not necessarily mean any estimate is inaccurate. Both can be correct for their corresponding parameter of interest.

3. Section 3. Page 4. It would be helpful to make it clearer that the model formulation (1), individual and joint likelihood contributions (2) and (3) correspond to the setting where there are a total of k studies, and each study contributes one T_i. The main technical results seem to focus on this setting (i.e., T_i's are independent) while the example in Section 2 corresponds to the setting where T_i's can be correlated.

4. Page 4, line 2 after model (1), typo in "which is true of some effect sizes"?

5. Section 3.2, page 6, line 49-50, the coefficients β_S is not necessarily a subset of the full vector of coefficients β, the full model with all covariates and a restricted model with a subset of covariates can both hold, the restricted model would correspond to the full model with the remaining variables marginalized out. β_S and the corresponding subset of β in the full model are not directly comparable.

6. Section 4, page 9, I fail to understand expression (11). Let T denote an estimator for a target parameter θ, then bias of an estimator is defined as $E(T) - \theta$. I don't understand the sentence "The bias term δ_{ij} refers to the bias induced in effect estimate i due to conditioning on missing pattern $\mathcal{R}_j$". In the current meta-regression setting, the effect estimate T_i are obtained irrespective of X_i. The potential bias due to missingness in X_i is only relevant when considering the coefficient of X_i in the meta-regression model relating T_i and X_i.

7. Section 4, page 9, please define the terms in equation (12). What is m_j? what are $f_{mj}(\cdot)$'s? With p covariates, there will be 2^p missing patterns. How would model (12) work? Are there additional assumptions required to model the missingness mechanisms?

8. Section 5.1, page 12, the w_i's in (18) and (19) are undefined.

9. Section 6. Please see my comment #2. In linear regression models, coefficients in a restricted model (i.e., based on a subset of covariates) are not necessarily a subset of the coefficients of the same variables in the full model with the presence of other covariates. The bias in shifting-case analyses should compare estimates of these coefficients to the true parameter values in the restricted model. More specifically, use linear regression models as a simple example:

$$
\begin{aligned}
E(Y \mid X_1, X_2) &= \alpha + \beta_1 X_1 + \beta_2 X_2 && (1)\\
E(Y \mid X_1) &= \alpha^* + \beta_1^* X_1 && (2)\\
E(Y \mid X_2) &= \tilde{\alpha} + \tilde{\beta}_2 X_2 && (3)
\end{aligned}
$$

Note that β_1^* in general is not equal to β_1, and $\tilde{\beta}_2$ is not equal to β_2. There are full distributions such that all models (1)-(3) are correct. When we fit models (2) and (3), the target parameters are β_1^* and $\tilde{\beta}_2$, not the original β_1 and β_2. They have different interpretations. The conditions the authors try to identify are those that yield $\beta_1^* = \beta_1$, and $\tilde{\beta}_2 = \beta_2$. When these conditions are not met, that does not mean $\hat{\beta}_S$ will be biased.