Review of: "IQOS® Cross-Sectional and Cohort US Study Documentation"

James Christopher (Jim) Anthony[1]

1 Michigan State University

Potential competing interests: Disclosure of potential conflict of interest: Dr. Hui G. Cheng studied with me and we published many papers together prior to 2019. Nothing on tobacco or nicotine products. I have had nothing to do with Phillip Morris USA or any of its subsidiaries, to the best of my knowledge.

In general, a clearly written study protocol. The figures and technical appendices are valuable additions to the text. Readers might wish to review those materials in advance of reading the article. The figures especially help to tell the story of what is planned.

The background is that the US Food & Drug Administration (FDA) is requiring Philip Morris USA to evaluate the effects of its claim about a retail Modified Risk Tobacco Product (MRTP) as described in the article. This protocol describes the firm's attempt to fulfill the FDA requirement.

The specification of ‘study population’ is tightly focused upon previously established MRTP consumers (focus on those who initiated MRTP use within six months prior to recruitment and initial assessment) and a comparison group of cigarette product smokers. The specification reads more like a description of a market research sample than a description of the population to which inferences are to be drawn.

It is possible that FDA provided these specifications and the firm is responding to the FDA stipulations. I wonder about the MRTP drawing from a broader consumer base that might include individuals who would not smoke traditional cigarettes, but who become users of the MRTP due to the claim that the product is a modified risk product. An analogy to consider is a cannabis user with lung problems who might not ‘smoke’ cannabis but who would consume an oral candy product with cannabinoid content. If I were again to be appointed to an FDA advisory committee, I would be asking about this form of MRTP effect, and whether the MRTP might be to the benefit of established MRTP users, but have externalities in the form of increasing the consumer base along these lines. The protocol does not seem to be oriented to investigating this possibility.

The rationale for truncating the age of participants at age 21 years should be clarified, and it might be wise to set an upper age limit to reduce sources of variation in response to the MRTP. Surely some MRTP consumers will be 18-20 years of age if not younger. Individuals younger than 21 might be drawn into the use of MRTP via the MRTP claim.

The definition of ‘established’ MRTP users is one that left-truncates the count distribution for number of times the MRTP has been used. This definition should ensure that the MRTP experience serves a reinforcing function for these consumers. Otherwise, the required number of MRTP experiences would not be reached. It would be great to have a

comparison of the number of individuals who try a conventional tobacco cigarette once and never again (which our research group is estimating) versus the number of individuals who try the MRTP once and never again (which the authors could estimate, but are not doing so in this protocol).

I express a little concern that the CrossSectional (CS) sample size might be too small to investigate everything the study team wishes to investigate, but the power analysis suggests a capacity to detect some effects of interest. (Latent variable analyses may be out of the question though.)

Even with some sample attrition over time, the Longitudinal Cohort (LC) sample size should be adequate. The difficult part will be to secure information about those who stop participation prematurely and to study the missingness mechanisms, whether they can be assumed to be MAR, MCAR, etc.

Why the firm would choose to hire a market research firm to recruit the sample is unclear, and the rigor of sampling and survey approaches used in market research should be monitored so that it is as good as NIH study sections require. Alternatives to market research might be valuable. Experienced NIH-caliber experts might be secured to help with quality control.

The analysis plan is sound but limited. As noted above, the CS sample size of 250 most likely will not be sufficient for latent variable modeling to help constrain noise in the measurements. The LC sample most likely will be sufficient, but there is a focus on population-averaged estimates and inferences (e.g., GEE), but no coverage of subject-specific alternatives that might yield smaller variances for estimates and inferences about subject-specific sources of variation that should be of interest to FDA.

If the participants are distributed across jurisdictions with variation in nicotine product taxes or retail sales prices, or with varying prevalences of tobacco smoking, a hierarchical model might be needed to take the local area variations into account along with the subject-specific sources of variation.

Finally, if there is a Data Safety Monitoring Board, I did not see one described. FDA might not require one, but there are advantages to a DSMB approach, if only to help constrain quality control lapses that can occur in longitudinal observational studies of this type.