UNIVERSITY OF CALIFORNIA SAN FRANCISCO

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

STANTON A. GLANTZ, PhD
Professor of Medicine (Cardiology)
Truth Initiative Distinguished Professor of Tobacco Control
Director, Center for Tobacco Control Research and Education

530 Parnassus Suite 366
San Francisco, CA 94143-1390
Phone: (415) 476-3893
Fax: (415) 514-9345
Stanton.Glantz@ucsf.edu

December 21, 2018

FDA Center for Tobacco Products
c/o Regulations.gov

RE: PMI's 6-month study "Evaluation of Biological and Functional Changes in Healthy Smokers After Switching to THS 2.2 for 26 Weeks (ZRHR-ERS-09 US) submitted in PMI IQOS MRTP June 8, 2018 amendment to FDA-2017-D-3001-0002 does not support claims of reduced risk.

PMI's MRTP application included their 3-month study of 24 non-cancer biomarkers of potential harm (which PMI calls "clinical risk endpoints," CRE) in humans using IQOS compared to conventional cigarettes. These biomarkers include measures of inflammation, oxidative stress, lipids, blood pressure, and lung function. (PMI did separate studies of biomarkers of exposure, several of which are carcinogens.) While PMI's application emphasizes that these biomarkers generally changed in positive directions, the data they submitted revealed no statistically detectable difference between IQOS and conventional cigarettes for 23 of the 24 BOPH in Americans and 10 of 13 in Japanese,[1] with the few significant differences were false positives. ***Thus, despite delivering lower levels of some toxicants, PMI's own data failed to show consistently lower risks of harm in humans using IQOS compared to conventional cigarettes.***

Their undated response, "The Difference between IQOS and Continued Smoking,"[2] presents two arguments against this conclusion:

- The original study submitted to FDA was "NOT DESIGNED to serve as the sole pivotal evidence with regards to CRE's and to show statistically significant changes in the CREs."
- PMI has a new, larger study 6-month human study comparing IQOS with conventional cigarettes that concluded that IQOS is less risky than conventional cigarettes.[3]

With regard to the first point, one is left with the question of why PMI submitted and represented data in the original MRTP application, when it now admits that the study was not designed to provide key evidence. They did not question the specific conclusions[1] that I drew about their 3-month study.

The new 6-month study (ZRHR-ERS-09 US) differs from the study presented in the original MRTP application in several important ways.

First, it is much larger (984 people in the new study compared to 79 in the US study and 112 in the Japanese study cited in the MRTP application). Making the study larger increases statistical power and makes it more likely to declare a difference statistically significant. This is a good thing.

Second, and of greater concern, the new study only considers 6 of the 24 non-cancer biomarkers in the earlier study, leaving the question of why PMI did not measure the other 18. (The 2 other biomarkers in the new study are biomarkers of exposure [CO and NNAL], which were not included in the earlier study and are not at issue in my paper.) Most of the things that they leave out are determined from blood tests, but they had to draw blood to measure the biomarkers they do report. The others are more detailed measures of lung function than the one reported in the new study and easily measured measures of blood pressure.

PMI should be expanding not dropping clinical endpoints because of evidence that IQOS is different from cigarettes.[4,5] For example, the data they presented in the MRTP application suggested that IQOS may be causing liver damage not observed in cigarettes.[6] Given the millions of dollars PMI's application represents, cost does not justify dropping these routine clinical measures. Their detailed presentation on the new study[3] does not address this question.

Third, PMI uses an arcane, little-used statistical method, the Hailperin-Rüger method, that was developed to *confirm* earlier studies.[7] (Neither I nor two biostatistics colleagues have seen this used in any recent clinical trials. A *PubMed* search with the keyword "Hailperin-Rüger" conducted on December 19, 2018, resulted in just one study.[8] The basic argument of the Hailperin-Rüger method is that it is overly cautious to require that all observed changes be statistically significant in order to *confirm* that a therapy works, and that if some lesser number of the variables change significantly, that should be good enough for a global test. The number of significant changes is specified in advance and the probability of a chance finding is adjusted.

PMI decided that if 5 of the 8 biomarkers (6 clinical risk and 2 exposure) changed in the direction of less risk, that would be enough to conclude that IQOS was less risky than conventional cigarettes. They do not provide a clear explanation of why they used 5, other than it was "more than half."

PMI justified using Hailperin-Rüger because "the probability of finding five significant tests ($p<0.05$) by chance alone is extremely low (0.006%)." This is a misleading statement because this low probability would only be the case a chance finding if *none* of the five variables actually changed. The probabilities are much higher when there are real changes.

So, in the new study, PMI went from considering changes in 24 clinical risk biomarkers in the original study to 8 in the new study to only requiring 5 to be statistically significant. That is a pretty major drop in the level of evidence PMI now suggests is sufficient to demonstrate that IQOS is less risky than cigarettes.

In the new study 5 of the changes were statistically significant, so PMI concluded that, overall, IQOS was better. Had they picked 6 in their plan, the overall results would not have been significant, even under the Hailperin-Rüger method's relaxed standards.

There are other problems with using the Hailperin-Rüger method. First, it is designed to *confirm* results of earlier studies. The earlier study did not convincingly show that IQOS was better than conventional cigarettes. Second, the usual way that Hailperin-Rüger is used is when you have several measures *of the same thing.* (For example, the one paper[8] located in PubMed that used Hailperin-Rüger assessed 10 different measures of neurological function and pre-specified that if 5 of the 10 were statistically significant, the global test would be considered statistically significant.)

The idea is that requiring all of them to change significantly is being too stringent a requirement to identify a change in lung function. In this case, PMI mixed apples and oranges by applying the text to a set of 6 clinical variables and 2 exposure that were measuring different underlying physiological processes.

PMI also used a one-tail test that assumes that one only need worry above improvements in the biomarkers without any concern for the possibility that they might worsen the biomarkers. (As noted above, PMI presented – but did not emphasize – other evidence in their MRTP application showing that IQOS caused problems not observed in cigarettes.[6]) I tell my students that, with rare exceptions, one should always do two-tail tests. Two-tail tests require larger differences to reach statistical significance, so by using a one-tail test, PMI made it easier to conclude changes were statistically significant. In this case the overall conclusion would have been the same with a two-tail test, so this bias did not make any practical difference, but they should have not used a one-tail test.

PMI's use of a one-tailed test was especially hypocritical since back in the early 1990's the tobacco companies sued the US EPA for using a one-tail test in their risk assessment that concluded that secondhand smoke caused lung cancer.[9] EPA used a one-tail test because they said it was inconceivable that secondhand smoke exposure would protect against lung cancer (the other tail). The irony there was that EPA would have reached the same conclusion using a two-tailed test.

All this raises the question of whether PMI manipulated the experimental design and analysis to get the desired conclusion, as they have done in the past.[10]

The law requires the MRTP applicant PMI to demonstrate, among other things, that IQOS, as it is actually used by consumers, will "significantly reduce harm and the risk of tobacco-related disease to individual tobacco users." Neither the original 3 month study nor the newer 6 month study meet this standard.

FDA should not rely on PMI's new study to support a conclusion that IQOS is less risky than conventional cigarettes.

Thank you for your consideration.

Stanton A. Glantz, PhD
Professor of Medicine
Truth Initiative Distinguished Professor in Tobacco Control
Director, Center for Tobacco Control Research and Education

References

1. Glantz S. PMI's Own in vivo Clinical Data on Biomarkers of Potential Harm in Americans Show that IQOS is Not Detectably Different from Conventional Cigarettes. *Tob Control.* 2018;27(Suppl 1):s9-s12. doi: 10.1136/tobaccocontrol-2018-054413. Epub 052018 Aug 054421.

2. Baker G, Harris C, Hankins M, et al. The Difference between IQOS and Continued Smoking. 2018; https://www.pmiscience.com/resources/docs/default-source/news-documents/the-difference-between-switching-to-iqos-and-continued-smokingef07ae852f88696a9e88ff050043f5e9.pdf. Accessed 19 Dec 2018.

3. PMI Research & Development. Study Results Overview: ZRHR-ERS-09 US (Evaluation of Biological and Functional Changes in Healthy Smokers After Switching to THS 2.2 for 26 Weeks. In PMI IQOS MRTP June 8, 2018 Amendment: Additional Information and Data from a Recently Completed Clinical Study (.zip – 1.3 GB) (added November 29, 2018). 2018; https://digitalmedia.hhs.gov/tobacco/static/mrtpa/PMP/June%208%2C%202018.zip. Accessed 18 Dec 2018.

4. Glantz S. Heated tobacco products: The example of IQOS. *Tobacco Control.* 2018;27(Suppl 1):s1-s6; DOI: 10.1136/tobaccocontrol-2018-054601.

5. St. Helen G, Jacob P, Nardone N, Benowitz N. IQOS: Examination of Philip Morris International's claim of reduced exposure. *Tob Control.* 2018;27(Suppl 1):s30-s36. doi: 10.1136/tobaccocontrol-2018-054321. Epub 052018 Aug 054329.

6. Chun L, Moazed F, Matthay M, Calfee C, Gotts J. Possible Hepatotoxicity of IQOS. *Tob Control.* 2018;27(Suppl 1):s39-s40. doi: 10.1136/tobaccocontrol-2018-054320. Epub 052018 Aug 054321.

7. Koch GG, Gansky SA. Statistical Considerations for Multiplicity in Confirmatory Protocols. *Drug Information Journal.* 1996;30(2):523-534.

8. Schellenberg R, Todorova A, Wedekind W, Schober F, Dimpfel W. Pathophysiology and psychopharmacology of dementia--a new study design. 2. Cyclandelate treatment--a placebo-controlled double-blind clinical trial. *Neuropsychobiology.* 1997;35(3):132-142.

9. Schachtman NA. EPA Post Hoc Statistical Tests – One Tail vs Two. 2012; http://schachtmanlaw.com/epa-post-hoc-statistical-tests-one-tail-vs-two/. Accessed 19 Dec 2018.

10. Wertz MS, Kyriss T, Paranjape S, Glantz SA. The toxic effects of cigarette additives. Philip Morris' project mix reconsidered: an analysis of documents released through litigation. *PLoS Medicine.* 2011;8(12):e1001145.