Cancun • Mexico
CORESTA
SSPT 2023

SMOKE SCIENCE and PRODUCT TECHNOLOGY Conference

SYMPOSIUM

Consumer Reported Outcome Measures: Science Starts with Measurement: Essential Measurement Science for Self-Report in Tobacco and Nicotine Product Research

11 October 2023

CORESTA

WEDNESDAY 11 OCTOBER 2023

Consumer Reported Outcome Measures (CROM) Symposium

Science starts with measurement: essential measurement science for self-report in tobacco and nicotine product research

Chair: Stacey McCAFFREY
Co-Chair: Mohamadi SARKAR

Starting at 14:20

No.	Time	Titles	Lead, Affiliation
CROM 01	20 min.	**Overview of the psychometric CROM guidelines**	Stacey McCAFFREY *Juul Labs, Inc.*
CROM 02	20 min.	**Overview of psychometrics: the science of measurement**	Stacey McCAFFREY *Juul Labs, Inc.*
BREAK – 5 minutes			
CROM Guest	20 min.	**Risk perceptions in tobacco regulatory science**	Alexander PERSOSKIE *FDA Center for Tobacco Products*
CROM 03	20 min.	**Evaluating the psychometric properties of a CROM for use with a different product for which it was developed: the modified e-cigarette evaluation questionnaire (MCEQ) as a case study**	Meghan MOREAN *Yale School of Medicine*
CROM 04	20 min.	**Further validation of the ABOUT-Dependence measure: Extending assessment of perceived dependence on tobacco and nicotine products to users of a heated tobacco product (*IQOS*)**	Esther AFOLALU *Philip Morris International*
CROM 05	20 min.	**An illustration that one size may not fit all: assessing invariance of the WISDM scale in PATH across youth and young adult cohorts**	Ryan BLACK *Juul Labs, Inc.*
CROM 06	20 min.	**Can individuals with limited health numeracy use quantitative scales to make ratings of risk perceptions?**	Saul SHIFFMAN *Pinney Associates*
BREAK – 10 minutes			
	20 min.	**Discussion and Q&A**	Mohamadi SARKAR *Altria Client Services* All speakers
Total: 175 minutes			

CROM 00 – Introduction

Science starts with measurement: essential measurement science for self-report in tobacco and nicotine product research

McCAFFREY(1); MOREAN M.(2); AFOLALU E.(3); BLACK R.(1); SHIFFMAN S.(4); SARKAR M.(5)

(1) Juul Labs Inc., 1000 F Street NW, Washington DC 20004, U.S.A.
(2) Yale School of Medicine, 333 Cedar Street, New Haven, CT 06510, U.S.A.
(3) Philip Morris Products S.A., Quai Jeanrenaud 5, CH-2000 Neuchatel, Switzerland
(4) Pinney Associates, Inc., 4800 Montgomery Ln #400, Bethesda, MD 20814, U.S.A.
(5) Altria Client Services LLC, 601 E Jackson St, Richmond, VA 23219, U.S.A.

Consumer reported outcome measures (CROM) are a critical component of tobacco regulatory science. Examples of CROM include measures of tobacco product dependence and risk perceptions. Like other fields of science, behavioral researchers require instruments (i.e., CROM) that are reliable and valid for measuring a given construct (e.g., risk perceptions), to ensure accuracy and replicability of measurement, and to support the validity of study conclusions. Researchers must also consider the context in which a CROM is applied, and whether a CROM valid for use in one context is valid in another. Such concerns are the focus of psychometrics, the science of behavioral measurement.

The CORESTA CROM Task Force (TF) is charged with establishing best practices and guidelines pertaining to the use of CROM in tobacco research. This symposium, organized by the CROM TF, will begin by introducing the TF's guidelines regarding the identification, development, validation, and implementation of CROM (McCaffrey). An overview of psychometrics will be provided and the importance of considering CROM psychometric functioning when used in a different context from which it was developed will be discussed (McCaffrey). To illustrate this point, two speakers (Morean and Afolalu) will present the application of CROMs to different products (e-cigarettes and IQOS) and the languages from which they were developed. Next, Black will demonstrate how invariance testing can be used to evaluate CROM functioning when used with different populations and discuss the consequences when such testing reveals non-invariance. Shiffman will present an approach for evaluating the appropriateness of numerical rating scales for individuals with limited numeracy to determine whether scores from such scales are comparable across those with limited vs. adequate numeracy. The discussant (Sarkar) will elaborate on the analogy between the methods used to establish inferences from clinical outcome measures compared to those used for self-reported measures.

CROM 01

Overview of the psychometric CROM guidelines

McCAFFREY S.

Juul Labs, Inc., Washington, DC, U.S.A.

This presentation will delve into the primary deliverable of the CORESTA Consumer Reported Outcome Measures (CROM) Working Group 02 (WG02) - guidelines and best practices with respect to Psychometric CROM for use in tobacco and nicotine product (TNP) research. After reviewing the definition of Psychometric CROM, noting differences between Psychometric and Descriptive CROM, the presentation will provide examples of Psychometric CROM commonly used in the TNP space. Next, WG02's consensus-based approach for guideline development will be described, which included literature review and consultation with subject matter experts. After providing context for the guidelines, including their purpose, scope, and intended audience, the presentation will cover the content of the guidelines. This content includes: an exercise for identifying the ideal characteristics of a Psychometric CROM within the context of the research study, recommendations for when and how to collect psychometric evidence to support a modified Psychometric CROM (based on the type and extent of CROM modifications), steps for developing and validating a new Psychometric CROM, and recommendations related to the implementation and interpretation of a Psychometric CROM. Finally, the presenter will share the status of the guidelines, including the publication/dissemination plan, and provide contact information for any audience members who are interested in learning more.

Dr. Stacey McCaffrey is a Psychometrician at JUUL Labs, Inc. where she leads the experimental behavioral research program. Dr. McCaffrey specializes in both qualitative and quantitative psychometric research methodologies, and has developed and validated measures of behavioral intentions, risk perception, claim comprehension, and respiratory symptoms for use in research as part of tobacco product applications to the FDA.

Prior to her work at JUUL Labs, Dr. McCaffrey was extensively involved in NIH and industry-funded research efforts targeting the opioid epidemic through Inflexxion, Inc. These efforts included development of a computer adaptive testing version of the Addiction Severity Index, as well as brief screening tools for adults who may be at risk for opioid misuse. As a consultant at PatientsLikeMe, Inc. she also led Robert Wood Johnson Foundation funded research to better understand patient priorities in healthcare, and developed several patient-reported outcome instruments to measure global and disease-specific health-related quality of life.

Dr. McCaffrey is leading the CORESTA consumer reported outcome measures (CROM) working group 02 (WG02) along with Esther Afolalu (PMI). She received her PhD in Clinical Psychology from Nova Southeastern University (Florida, USA) in 2015. Dr. McCaffrey is also a licensed clinical psychologist.

CROM 02

Overview of psychometrics: the science of measurement

McCAFFREY S.

Juul Labs, Inc., Washington, DC, U.S.A.

Consistent with other areas of science, behavioral science requires objective measurement that is precise, replicable, and measures what it is intended to measure. When measuring behavioral constructs such as dependence or risk perceptions, it is critical that the researcher has reliable and valid tools. Psychometrics is the field of science concerned with the psychometric functioning (reliability, validity) of these self-report measuring tools, which we refer to as "consumer reported outcome measures" (CROM).

The Psychometric CROM guidelines developed by CORESTA CROM Working Group 02 (WG02) present best practices for the selection, development and validation, modification, and implementation of Psychometric CROM for use in research on tobacco and nicotine products (TNPs). It is important that a behavioral researcher working in the area of TNPs has basic knowledge of foundational psychometric concepts in order to appreciate and utilize these guidelines most effectively. This presentation will review basic psychometric properties, such as types of reliability and validity, and present examples of both qualitative and quantitative methodologies for evaluating psychometric functioning of a new, modified, or existing CROM. For example, cognitive debriefing interviews will be reviewed as a qualitative strategy to evaluate and improve content validity of a CROM, and confirmatory factor analysis will be discussed as a quantitative strategy to evaluate measurement invariance across product categories or populations. Additionally, the presentation will cover psychometric considerations when identifying an appropriate CROM, such as evaluating the match between an existing CROM's psychometric functioning within a particular context of use and the psychometric properties of greatest importance within the context of the study. Finally, psychometric considerations when implementing a single or multiple CROM (i.e., when multiple CROM are combined into a survey) in a research study will be discussed.

CROM 03

Evaluating the psychometric properties of a CROM for use with a different product for which it was developed: the modified e-cigarette evaluation questionnaire (MCEQ) as a case study

MOREAN M.

Yale School of Medicine, 333 Cedar Street, New Haven, CT 06510, U.S.A.

The Cigarette Evaluation Questionnaire, and its revised version, the Modified Cigarette Evaluation Questionnaire (MCEQ), have been used for decades to assess the subjective reinforcing and aversive effects of cigarette smoking. While not all measures are good candidates for translation for use with a different product(s), the MCEQ's items appeared relevant to e-cigarette use. Thus, we aimed to examine the psychometric properties of the Modified E-cigarette Evaluation Questionnaire (MECEQ).

In Summer 2021, 857 adults completed an anonymous online survey (52.4 % male; 40.84 [12.25] years old; 62.8 % non-Hispanic white; 22.4 % daily e-cigarette use). Analyses included exploratory/confirmatory factor analyses to confirm the original structure and/or identify alternate latent structure(s), internal consistency, measurement invariance, between-group differences, and relationships with vaping outcomes.

The results showed that the original five-factor structure and a novel four-factor structure were supported. All multi-item subscales were internally consistent. Both the five-factor and four-factor versions reached scalar invariance across multiple participant subgroups, could detect between-groups differences, and were associated with past-month vaping frequency and dependence.

The results strengthen the interpretability of previously published work using the five-factor structure and provide an alternative, psychometrically-sound, four-factor scoring approach. Future research is needed to evaluate invariance between the MCEQ and MECEQ before subjective effects of smoking and vaping can be compared directly.

Dr. Meghan Morean completed her undergraduate training at Brown University (2004), her PhD in clinical psychology from Yale University (2011), and her postdoctoral fellowship at the Yale School of Medicine (2011-2014). She then was employed at Oberlin College (2014-2020), earning tenure as an associate professor. While at Oberlin, she maintained her affiliation with Yale (adjunct assistant professor). In Summer 2020 she returned to the Yale University School of Medicine. Dr. Morean's program of research focuses on youth and adult use of nicotine/tobacco products, cannabis, and alcohol. She has expertise in measurement development and has contributed numerous substance-relevant measures to the field.

CROM 04

Further validation of the ABOUT-Dependence measure: Extending assessment of perceived dependence on tobacco and nicotine products to users of a heated tobacco product (*IQOS*)

AFOLALU E.

Philip Morris Products S.A. (part of Philip Morris International), Quai Jeanrenaud 5, Neuchâtel, Switzerland

The ABOUT™–Dependence is a Consumer Reported Outcomes Measure (CROM) developed to assess perceived dependence associated with tobacco and/or nicotine product (TNP) use. It was developed and validated in US English for TNP users who were either exclusive or poly users of different TNPs (including cigarettes, e-cigarettes, cigars/cigarillos, smokeless tobacco, pipes, and waterpipes) and nicotine replacement therapy (NRTs). The initial validation confirmed the psychometric performance of a 12-item version of the CROM consisting of three domains: Extent of Use (2 items); Signs and Symptoms (5 items); and Behavioral Impact (5 items).

This presentation will describe additional research conducted to assess the validity of the CROM when applied to heated tobacco products (HTPs) use (specifically, *IQOS*) and to evaluate its cross-cultural validity. The research included analyses of data collected as part of an online survey with 1320 adult TNP users in the US, Germany, Italy, Russia, and Japan who were cigarette smokers or who had switched from smoking cigarettes to using either HTPs (*IQOS*) or other TNPs (e.g., e-cigarettes) exclusively or were dual users of cigarettes and another TNP. The presentation will outline the psychometric techniques (Rasch Measurement Theory (RMT) and Classical Test Theory (CTT)) used to confirm the measurement properties of the CROM in an extended frame of reference. The findings largely aligned with the previous validation of the CROM and provided supportive evidence of the CROM's reliability and validity for the whole sample (N = 1320) and a subset of exclusive *IQOS* users (N = 263). The instrument's stability across different countries/languages was also acceptable, although differential item functioning (DIF) by country/language was observed for some items.

Overall, the research further supports the ABOUT™–Dependence as a psychometrically valid measure of perceived dependence with broad cross-cultural applicability to different types of TNPs and product use behaviors.

Esther F. Afolalu, PhD, is a Senior Behavioral Scientist at Philips Morris International, where she provides technical leadership for behavioral research projects, including the development, validation, and implementation of the ABOUT™ Toolbox portfolio of Consumer Reported Outcomes Measures) (CROMs) for assessment of consumer perception and behaviors related to the use of tobacco and nicotine products. She is additionally involved in the CORESTA CROM Task Force as a co-coordinator of the Psychometric CROM Working Group.

Her research expertise spans the fields of Behavioral Science, Clinical and Health Psychology, Population Health, Outcomes Research, and Epidemiology. She holds a PhD in Psychology (University of Warwick, UK) and her doctoral research focused on multi-methodological experimental and observational epidemiological studies on the associations of sleep disturbances with pain and health outcomes. She also has past experience in conducting pharmaceutical clinical trials, specifically in areas such as psychopharmacology.

CROM 05

An illustration that one size may not fit all: assessing invariance of the WISDM scale in PATH across youth and young adult cohorts

BLACK R.A.(1); SHIFFMAN S.(2); HANNON M.J.(2)

(1) Juul Labs, Inc., Washington, DC, U.S.A
(2) Pinney Associates, Inc., Pittsburgh, PA, U.S.A.

Measuring dependence on ENDS in youth and adults is of interest, as dependence bears on continued use of such products. Dependence is an abstract, complex concept whose measurement can be highly context-specific. Typically, different scales are used to measure dependence in youth and adults. For example, 6 items from the adult Wisconsin Inventory of Smoking Dependence Motives (WISDM) scale are used to measure dependence in the PATH youth survey, while the 16-item Tobacco Dependence (TD) Index, which also includes some WISDM items, is used in the PATH adult survey. Some researchers have assessed dependence on ENDS using a single scale in a combined sample of youth and adults, without confirming that the scale is *functioning* equivalently across the age cohorts. As an illustration, we used psychometric analyses to assess the equivalence of the WISDM scale in a sample of PATH youth (15-17 years) and young adults (18-34 years). Factorial invariance of the WISDM scale was assessed across youth and young adult cohorts via confirmatory factor analysis (CFA). Data were obtained from youth (N=229) and young adult (N=761) respondents to PATH Wave 5 Youth and Adult surveys who reported current (past 30-day), exclusive ENDS use. In CFA-based invariance testing, the WISDM scale did not function equally across age cohorts, indicating that comparing or combining youth and young adult dependence scores may be problematic. The findings reinforce that one cannot *assume* that a seemingly identical scale can be used to measure dependence on ENDS in different age cohorts. A proper psychometric analysis should be undertaken to assess whether the scale is functioning equivalently. Not doing so risks drawing inaccurate conclusions regarding levels of dependence, which could lead to ill-advised public health policy.

Dr. Ryan Black is a clinical and research psychologist with expertise in psychometrics, statistical methods and population health impact modeling. He currently works in the regulatory science department at Juul Labs, Inc., serving as VP of Clinical and Population Sciences. Dr. Black previously led the population science team in the regulatory science department at Altria. As part of his role, he led the PHIM work and presented the ALCS PHIM at various venues, including TPSAC, CORESTA, FDLI and academic institutions. Dr. Black has held the position of Director of Biostatistics and Methodology at Inflexxion, Inc., in which he served as co-PI/lead biostatistician on several NIH-funded behavioral research grants. During this time, Dr. Black developed and executed the data analytic strategy used to evaluate abuse of prescription opioid products as part of broader postmarket surveillance programs in support of regulatory engagement by various prescription opioid manufacturers. Dr. Black has published extensively in the areas of psychometric methods, substance use and tobacco regulatory research.

CROM 06

Can individuals with limited health numeracy use quantitative scales to make ratings of risk perceptions?

SHIFFMAN S.(1); HANNON M.J.(1); McCAFFREY S.(2)

(1) Pinney Associates, Inc., Pittsburgh, PA, U.S.A.
(2) Juul Labs, Inc., Washington, DC, U.S.A.

Assessment of risk perceptions of lower-risk nicotine products such as electronic nicotine delivery systems (ENDS) is an important area of regulatory research. Psychometric challenges emerge when attempting to estimate and quantify their risk perceptions. The ability of individuals with limited numeracy skills to use numerical (quantitative) rating scales has been questioned. We report an online study using numerical ratings of risk perception obtained from individuals with varied numeracy. Participants were 12,557 adults including smokers, dual users, former tobacco users and never-users who viewed information about JUUL, and who were randomized to see or not see a reduced-exposure message. (Neither tobacco use status nor message exposure are part of the psychometric analyses.) Participants completed the Newest Vital Sign (NVS), which asks respondents to interpret a Food Facts label containing quantitative information, and requiring quantitative reasoning and computation. The NVS has been used to assess numeracy as well as literacy. The NVS classified 29 % of participants as having limited health numeracy ("LHN"). Using numerical scales from 0 % to 100 % harmful to health (in 10 % increments), participants rated the overall risk of harm from using JUUL, and the likelihood (0-100 %) of suffering four specific diseases (mouth cancer, lung cancer, heart diseases, respiratory disease). They also rated the risk of using JUUL using a commonly-used ordinal qualitative (descriptive) response scale (4-point scale from not at all harmful to very harmful). The psychometric properties of the numerical ratings were similar for LHN as for those with adequate health literacy (AHN). The numerical ratings of JUUL risk in the two groups showed a nearly identical orderly relationship to the descriptive perceived risk ratings, with very similar means at each qualitative level. In analyses considering the disease-specific ratings as a scale, results demonstrated equivalence between ratings from LHN and AHN individuals. The correlations among ratings of the four diseases were nearly identical for LHN as for AHN participants. Moreover, in analyses based on confirmatory factor analysis, both groups showed robust fit to a one-factor model. Tests of invariance between LHN and AHN found the scale to demonstrate configural, metric, and scalar invariance between AHN and LHN. Thus, analyses show that individuals with limited numeracy skills were able to make meaningful ratings using numeric scales, comparable to those obtained from respondents with adequate numeracy.

Saul Shiffman, Ph.D. serves as Senior Scientific Advisor at Pinney Associates, which consults to JUUL Laboratories on e-cigarettes and tobacco harm reduction. He is also Emeritus Professor of Psychology (Clinical and Health Psychology), Psychiatry, Pharmaceutical Sciences, and Clinical Translational Science at the University of Pittsburgh.

Dr. Shiffman has published over 450 scientific papers on topics including smoking patterns, nicotine dependence, smoking cessation and relapse, smoking cessation treatment, e-cigarette use, and tobacco harm reduction, as well as on measurement and research methods. His papers have received over 50,000 citations in the scientific literature. Dr. Shiffman has received a number of awards, including the Society for Research on Nicotine and Tobacco's award for "ground-breaking advances in clinical research."

Dr. Shiffman has been involved in a number of regulatory submissions to the US Food and Drug Administration's Center for Tobacco Products, and presented at the Tobacco Products Scientific Advisory Committee, as well as other FDA Advisory Committees.

CROM Guest

Risk perceptions in tobacco regulatory science

PERSOSKIE A.(1); O'BRIEN E.(2)

(1) Supervisory Social Scientist, Division of Population Health Science, Office of Science, FDA Center for Tobacco Products, U.S.A.
(2) Supervisory Social Scientist, Division of Research and Evaluation, Office of Health Communication and Education, FDA Center for Tobacco Products, U.S.A.