The PhenX Toolkit: measures for tobacco regulatory research

Gary E Swan,[1] Tabitha P Hendershot,[2] Carol M Hamilton,[2] Dana M van Bemmel,[3] Kay L Wanke ,[4] PhenX Tobacco Regulatory Research Panel

ABSTRACT
The use of standard measures and common data elements in the conduct of tobacco regulatory research (TRR) will improve data comparability and facilitate cross-study analyses and replication of findings to inform the regulatory decisions and actions of the US Food and Drug Administration (FDA) Center for Tobacco Products (CTP). This overview paper introduces the PhenX (consensus measures for Phenotypes and eXposures) initiative to identify common measures and common data elements for use in TRR which have been disseminated in the PhenX Toolkit (https://www.phenxtoolkit.org/). It describes the work of the scientific panel that provided guidance to the project and the TRR Core Collection of measures recommended for use by investigators conducting human-subject tobacco regulatory or related tobacco control research. The use of PhenX TRR measures will promote collaborative research and facilitate data interoperability and validation of outcome assessments to inform the FDA CTP's regulatory activities.

Tobacco regulatory research (TRR) is a nascent but fast-growing area of scientific inquiry due to the passage of the Family Smoking Prevention and Tobacco Control Act (TCA) of 2009.[1] As a result of the TCA, the Food and Drug Administration (FDA) established the Center for Tobacco Products (CTP); its mission is to protect Americans from tobacco-related death and disease through regulation of the manufacture, distribution and marketing of tobacco products. CTP relies on a sound science base to support its regulatory activities.[2] Since 2011, the CTP Office of Science has partnered with the National Institutes of Health (NIH) through the NIH coordination office, the Tobacco Regulatory Science Program (TRSP), to develop a tobacco regulatory research portfolio that builds on the NIH's decades-long investment in tobacco-related research.[3] One way to leverage the investment in tobacco regulatory science is the establishment of a standard set of measures that can be used across TRR studies to improve data compatibility. The ability to combine data across multiple studies creates larger sample sizes to support more powerful and replicable analyses, ultimately aiding in validation of findings and increasing the scientific impact of individual studies. As NIH and CTP share the goal of advancing science by improving the yield and impact of their research portfolios, they partnered to fund an administrative supplement to the PhenX (consensus measures for Phenotypes and eXposures) initiative to identify a set of measures to promote data comparability across TRR studies.

This *Tobacco Control* supplement presents the recommended measures that can serve as common data elements in TRR. The TRR measures, disseminated through the PhenX Toolkit (https://www.phenxtoolkit.org), provide the research community with a common currency of standard measures for data collection. The measures were selected by working groups (WGs) of domain experts who recommended well-established, broadly validated measurements within scientific topic areas. The broader scientific community was then invited to comment on the selections before the final recommendations were made for inclusion in the PhenX Toolkit. The TRR Collections build on the existing catalogue of measures within the PhenX Toolkit, which now include a total of 730 measurement protocols across 25 broad research domains and 5 specialty research areas: Tobacco Regulatory Research, Substance Abuse and Addiction Research, Mental Health Research, Sickle Cell Disease Research, and Hemophilia Inhibitor Research. As such, researchers can find protocols in the Toolkit to address a variety of research areas beyond TRR, from anthropometrics to clinical disease outcomes. In addition to the measurement protocols, the Toolkit provides numerous resources—such as data collection worksheets, data dictionaries, general references and translations (when available)—all of which make it easy for researchers to add one or more PhenX measures to their data collection forms. Already accessed by investigators across the USA and the world, and recommended for use in approximately 325 NIH Funding Opportunity Announcements, PhenX measures facilitate the collection of data in a consistent format using established measurement protocols.

The papers in this special issue include: (1) this overview paper, which provides the vision and goals for the project, an overview of the PhenX TRR consensus process and a description of the TRR Core measures; (2) five papers outlining the host, agent, vector and environment (HAVE) domain-specific collections of measures, the process for coming to consensus and the measurement and research gaps in these fields; (3) a portfolio analysis of NIH electronic cigarette (e-cigarette) grants funded by NIH and FDA, analysing the balance of proposed measures across the domain framework of the TRR Collections within a targeted TRR topic and (4) a systematic review of measures in risk perception, one of the gaps identified during the project.

OVERVIEW OF THE PHENX TRR PROJECT

Having recognised the need for standard measurement protocols and the impact their adoption could have on genome-wide association studies (GWAS), the NIH National Human Genome Research Institute first funded RTI International to develop PhenX in 2007. PhenX uses a consensus process,[4] assembling groups of experts to identify measurement protocols for the PhenX Toolkit, using the criteria shown in box 1. It is a selective, comprehensive catalogue of data collection protocols curated by domain experts who identify a prioritised set of consensus protocols.

THE TRR PANEL

In 2013, representatives from the NIH TRSP and FDA CTP convened a TRR Panel (TRRP) to facilitate the development of a collection of TRR measurement protocols for release into the PhenX Toolkit. The TRRP comprises nine academic and federal scientists, each of whom has an established research portfolio and expertise in the field of TRR. The two cochairs

[1]Stanford Prevention Research Center, Stanford University School of Medicine, Stanford, California, USA
[2]RTI International, Research Triangle Park, North Carolina, USA
[3]Center for Tobacco Products, Silver Spring, Maryland, USA
[4]Office of Disease Prevention, National Institutes of Health, Bethesda, Maryland, USA

Correspondence to Dr Gary E Swan, Stanford Prevention Research Center, Stanford University School of Medicine, Stanford, CA 94305, USA; gswan@stanford.edu

Box 1 Criteria used for selecting PhenX measures

The measure should be:
- Clearly defined.
- Well-established.
- Broadly applicable.
- Validated.
- Reproducible.
- Specific.
- Reliable.

Additional criteria for selecting the measures include the following:
- Acceptable burden to participants and investigators.
- Cross-cutting relevance for populations groups and for diseases and conditions.
- Open-source software (if required) preferred.
- Brevity.
- Acceptance by the research community
- Existing standard measurement protocols.

The final set of measures should cover the scope of the domain.

Table 1 TRR Core: Tier 1 measures in the PhenX Toolkit

Measure name (PhenX ID#)	Domain/Collection
Current Age (10100)	Demographics
Ethnicity (10500)	Demographics
Race (10600)	Demographics
Gender (10700)	Demographics
Current Marital Status (10900)	Demographics
Current Educational Attainment (11000)	Demographics
Current Employment Status (11300)	Demographics
Use of Tobacco Products (741400)	TRR: Vector
Cigarette Nicotine Dependence (31000)	Alcohol, Tobacco and Other Substances
Electronic Nicotine Product Device Type (760100)	TRR Core: Tier 1

PhenX, consensus measures for **Phen**otypes and e**X**posures; TRR, Tobacco Regulatory Research.

of the TRRP led the consensus process and provided guidance for the selection of TRR Core measures.

TRRP: DETERMINING THE SCOPE OF THE TRR COLLECTIONS

The TRRP defined the scope of the overall project and determined the focus of each of the TRR WGs. The TRRP adopted the HAVE model as the organising framework for the TRR Collections because of its use as the conceptual model for tobacco surveillance and evaluation by the 2002 National Tobacco Monitoring, Research and Evaluation Workshop.[5] The HAVE model is adapted from the classical infectious diseases paradigm, where the Host is the biological organism (here, the tobacco product user or potential user) affected by the toxic or infectious Agent that causes disease or harm (here, the tobacco product), delivered by the Vector (here, the tobacco product manufacturers, marketers and sellers) within the context of the Environment (here, the environment includes social, cultural, economic, public health and policy factors).[6 7] The TRRP identified experts for each of the HAVE domains, convening two Host WGs due to the large number of Host measurement elements, for the following five WGs:

- **Host: Social/Cognitive**: interpersonal and intrapersonal factors influencing tobacco use
- **Host: Biobehavioral**: behavioural and biological factors, including product use and exposure
- **Agent**: characteristics of tobacco products
- **Vector**: industry and retailer activities
- **Environment**: environmental influences on tobacco use

TRRP: DEVELOPMENT OF THE TRR CORE COLLECTIONS

In addition to being responsible for overseeing the work of the TRR WGs, the TRRP was charged with establishing a collection of core measures to be used across TRR studies. The Core Collection includes two tiers of measures for use by TRR investigators. Core: Tier 1 measures were deemed relevant and essential to all areas of TRR and require very little time to collect (table 1). Core: Tier 2 measures are relevant to many areas of TRR but are more specialised and may require more time to administer (table 2).

The Core measures provide a minimum set of common data elements to be used across TRR, enhancing the opportunity for cross-study collaboration.[8] The TRRP used the consensus process to develop the Core Collection. The TRRP initiated the process by reviewing complementary measures already in the Toolkit to identify potential areas for new content and then considering their relation to both TRR and the already-established Core Collections for Mental Health Research (https://www.phenxtoolkit.org/collections/view/1) and Substance Abuse and Addiction Research (https://www.phenxtoolkit.org/collections/view/2), two scientific domains relevant to the conduct of TRR. The TRRP had a vigorous discussion regarding prioritisation of demographic versus TRR-centric measures in Core Tier 1. On the one hand, there was interest in populating this Core with measures that would be used specifically for TRR studies. On the other hand, demographic measures help characterise a study population and facilitate comparison with other studies, both within and outside of the TRR community. Following establishment of the TRRP recommendations for Core measures, PhenX conducted outreach as part of its consensus process,[4] seeking feedback from the scientific community between 16 May and 3 June 2016. Thirty-five people responded to Community Outreach emails; 27 provided ratings on the value of the various measures recommended. Overall, respondents rated the proposed measures as being valuable to TRR.

After consideration of the feedback collected from the Community Outreach, the TRRP selected 10 measures for Core Tier 1 (table 1) and 14 measures for Core Tier 2 (table 2) for inclusion in the Toolkit. These were subsequently approved by the PhenX Steering Committee. The TRR Core Collections were released into the Toolkit on 30 August 2016, V.16.0. Tables 1 and 2 present brief descriptions of these measures along with PhenX identification numbers to facilitate reader access to each recommended measure on the Toolkit website. The website provides protocols and scoring procedures for each measure, identifies the populations and tobacco products for which it has been validated for use (if applicable) and identifies the source (eg, national survey, published manuscript). Due to similarities between tobacco regulatory and tobacco control research, some measures included in the TRR are outside the regulatory authorities of the TCA. Although

Table 2 TRR Core: Tier 2 measures in the PhenX Toolkit

Measure name (PhenX ID#)	Domain/Collection
Annual Family Income (11100)	Demographics
Child-reported Parental Education Attainment (210200)	Social Environments
Self-reported General Health Status (770100)	TRR Core: Tier 2
Veteran Status (770300)	TRR Core: Tier 2
Peer and Family Influence on Smoking (710900)	TRR Host: Social/Cognitive
Flavor preference- e-cigarettes (720600) Protocols for Adults and Adolescents	TRR Host: Biobehavioral
Tobacco Brand and Variety (730700) Protocols for Cigarettes, Cigars and Smokeless Tobacco	TRR: Agent
Self-reported Tobacco Product Price Paid (740800) Protocols for Cigarettes, Cigars, Smokeless Tobacco and e-cigarettes	TRR: Vector
Exposure to Tobacco Advertising on the Internet and Social Media (740100)	TRR: Vector
Internalising, Externalising and Substance Use Disorders Screener (580100)	Substance Abuse and Addiction Core: Tier 2
Self-reported Craving (520300) Protocols for General Craving, and for Tobacco, Alcohol, Cocaine, Marijuana and Heroin	Substance Abuse and Addiction: Substance-specific Intermediate Phenotypes
Smoking Quit Attempts (71000)	Cancer
Passive Smoke Exposure (70300)	Cancer
Blunt Use (770200)	TRR Core: Tier 2

TRR, Tobacco Regulatory Research.

implementation of policies such as state and local tobacco policies are outside TCA regulation, data from those policies may inform tobacco regulatory research and regulation.

DISCUSSION

The Core Collection and five Specialty Collections provide a set of recommended consensus measures to serve as common data elements for TRR studies that build on the existing catalogue of PhenX measures. This special issue of *Tobacco Control* includes five papers that summarise the process and outcomes of each of the WGs: (1) Host: Social/Cognitive,[9] (2) Host: Biobehavioral,[10] (3) Agent,[11] (4) Vector[12] and (5) Environment.[13] The remaining two papers provide the results of NIH programme staff investigations of issues raised during the PhenX TRR deliberations: an analysis of the HAVE measurement domains in funded grant applications investigating e-cigarettes,[14] and a review of the literature in measures of product-specific risk perceptions, a major gap identified during the TRR PhenX process.[15]

One of the consistent conclusions across WGs regarding the state-of-the-art of tobacco regulatory science is the need for replication and validation of measures developed for use with cigarettes or cigarette smokers that are now being adapted for measurement of other tobacco products. There are many reasons why translating directly from cigarette use to other tobacco product use may not be straightforward, including the nature of the newer products, the formulation of their constituents and the interaction between the user and the product itself. For instance, a measure of chronic use of cigarettes may be qualitatively different from that of chronic use of e-cigarettes. As these products continue to evolve, sustained effort is needed to validate these adapted measures and re-evaluate the validity of long-established measures. For example, the field may need to reconceptualise basic assumptions about dependence, given the impact of dual and polytobacco product use as well as other substance use. The introduction of novel, non-combustible tobacco products into the marketplace along with conventional cigarettes increases the likelihood that one or more of these products will be used to supplement or heighten the addictive potential of tobacco product use. Naturalistic studies of polyproduct use will inform TRR measurement efforts, including assessment of the frequency, duration and intensity of each of the products used. Measurement challenges also arise with tobacco products that are used in conjunction with other substances such as marijuana. The extent to which e-cigarettes or cigars, for instance, are modified to deliver substances other than nicotine needs to be assessed. A sustained effort will be needed to support ongoing renewal of measurement methods applied to existing and new tobacco product use.

In the process of selecting measures to include in the PhenX Toolkit, the WGs noted the lack of availability of many well-established HAVE measures for adolescents. Newly introduced products often rely on direct-to-consumer marketing channels such as social media and web-based advertising, which appeal to adolescents. Many of these marketing-exposure pathways are obvious, while others may not be as apparent—such as algorithm-based advertising to target and drive tobacco product content to adolescents and their friends based on their likes and dislikes. The field will most certainly be challenged by the question of how best to measure the effect of tobacco product marketing through social media platforms on adolescent product use. The adoption of the recommended PhenX measures by TRR investigators across varying populations and product categories will result in data sets large enough to support serious investigations of these influences.

An important observation resulting from the review of available measures was the disproportionate number relevant to the tobacco product user. The analysis of the portfolio of funded NIH e-cigarette research applications by Garcia-Cazarin and colleagues confirmed a predominant use of measures from the Host domains almost to the exclusion of measures from the other domains.[14] This suggests that future grant applicants should consider including measures of the Agent, Vector

and Environment. Doing so will provide a more comprehensive picture of the landscape influencing tobacco product users. Having measures from each of the domains within the same study will support an examination of the extent to which host characteristics interact with Agent, vVector and Environment factors to influence tobacco product use and will maintain fidelity to the overarching theoretical model. A more comprehensive coverage of HAVE components may also provide additional findings to inform regulatory activities.

Measurement of risk perception was identified by the Host: Social/Cognitive WG as a priority measurement construct; yet the WG deliberations concluded that a widely accepted standard was not available. As such, Kaufman, the WG's NIH liaison, and colleagues conducted a review to characterise the risk perception measures used in tobacco control research and to evaluate the degree to which recent research incorporated best practices advised by experts in the field. They found little consistency across studies in risk perception measures, concluding that advances in risk perception measurement and adoption of recommended standards are key to advances in the science. This finding is consistent with the NIH portfolio analysis conducted by Conway and colleagues that found a lack of measure commonality across funded research proposals in substance abuse and addiction (SAA) sciences, which provided the impetus for establishing the SAA Core and Specialty Collections within the PhenX Toolkit.[16]

The TRR Collection in the PhenX Toolkit provides a set of expert-recommended, prioritised measures as a resource for investigators conducting TRR. By providing investigators with a set of tools and resources to facilitate sharing, comparing, replicating findings and integrating data from multiple sources, the NIH TRSP and FDA CTP seek to advance science by improving the yield and impact of TRR.[8] These benefits include: (1) easy access to a freely available catalogue of measures for a wide variety of constructs and variables; (2) a reference for research applications and publications to the PhenX imprimatur, founded on expert recommendations and scientific consensus; (3) enrichment of research studies through the inclusion of easily administered, low-burden measures for variables and constructs that enable future hypothesis testing and (4) opportunities for future collaborative studies and/or replication of findings; collectively enhancing TRR while advancing a culture of scientific collaboration.

Acknowledgements We gratefully acknowledge the contributions of the anonymous members of the scientific community who commented on the measures proposed for inclusion. We also acknowledge the contribution of Deborah Maiese, RTI International, who led the recruitment of TRRP members and the orientation to the PhenX process, and the contributions of the RTI editorial team, including Michelle Bogus, Judy Cannada, August Gering, Loraine Monroe and Amy Morrow for providing expert editorial review of the manuscript.

Collaborators GROUP AUTHORSHIP STATEMENT The following are members of the PhenX Tobacco Regulatory Research Panel (TRRP): Cochairs Gary E Swan, Stanford Prevention Research Center and Judith (Jodi) J Prochaska, Stanford University; TRRP members Neal L Benowitz, University of California, San Francisco, Kevin P Conway, National Institute on Drug Abuse (currently with RTI International), Andrew Hyland, Roswell Park Comprehensive Cancer Center, Robin J Mermelstein, University of Illinois at Chicago, Dana M van Bemmel, US Food and Drug Administration Center for Tobacco Products, Kay L Wanke, National Institutes of Health Office of Disease Prevention and Gordon B Willis, National Cancer Institute; and the PhenX team from RTI International, including Principle Investigator Carol M Hamilton and TRR Project Lead Tabitha Hendershot.

Contributors GES, TPH, DMvB and KLW led the conceptualisation and writing of the manuscript. CMH reviewed the draft and provided substantive content, comments and revisions. The cochairs and members of the TRRP identified the overarching scope of the TRR project, nominated and approved WG experts, identified preliminary measurement constructs for WG consideration, participated in the WGs as liaisons and approved the HAVE Specialty Collections proposed by the WGs. The cochairs and members of the TRRP proposed preliminary measures and voted on final measures included in the PhenX TRR Core Collections. NIH Project Coordinator KLW proposed the PhenX TRR initiative and contributed to its execution and completion. The RTI PhenX team developed, coordinated and facilitated the TRRP and WG process, including project oversight and leadership (TPH, CMH) and project management Darigg C. Brown. PhenX NIH Program Official Erin M Ramos, National Human Genome Research Institute, provided project guidance and funding coordination.

Funding Activities reported in this publication were supported by grant number U41HG007050 from the National Human Genome Research Institute (NHGRI) and U41HG007050-01S1 from the NHGRI and US Food and Drug Administration (FDA) Center for Tobacco Products.

Disclaimer The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or the FDA.

Competing interests We have read and understood the Tobacco Control policy on the declaration of interests and declare that we have no competing interests.

Provenance and peer review Not commissioned; internally peer reviewed.

To cite Swan GE, Hendershot TP, Hamilton CM, *et al*. *Tob Control* 2020;**29**:s1–s4.

Tob Control 2020;**29**:s1–s4.
doi:10.1136/tobaccocontrol-2019-055125

ORCID iD
Kay L Wanke http://orcid.org/0000-0003-0660-9522

REFERENCES

1 US Department of Health and Human Services, Food and Drug Administration. Family smoking prevention and tobacco control act. Public law 111-31, 22. Available: https://www.fda.gov/TobaccoProducts/Labeling/RulesRegulationsGuidance/ucm262084.htm [Accessed 9 Apr 2019].
2 Ashley DL, Backinger CL. The food and drug Administration's regulation of tobacco. *American Journal of Preventive Medicine* 2012;43:S255–S263.
3 Backinger CL, Meissner HI, Ashley DL. The FDA "Deeming Rule" and Tobacco Regulatory Research. *Tob Regul Sci* 2016;2:290–3.
4 Maiese DR, Hendershot TP, Strader LC, *et al*. *PhenX—Establishing a consensus process to select common measures for collaborative Research (RTI press publication No. MR-0027-1310*. Research Triangle Park, NC: RTI Press, 2013.
5 Giovino GA, Biener L, Hartman AM, *et al*. Monitoring the tobacco use epidemic I. overview: optimizing measurement to facilitate change. *Prev Med* 2009;48:S4–S10.
6 Orleans CT, Slade J, Addiction N. *Principles and management*. New York: Oxford University Press, 1993.
7 Giovino GA. Epidemiology of tobacco use in the United States. *Oncogene* 2002;21:7326–40.
8 National Institutes of Health, U.S. Food and Drug Administration, Center for Tobacco Products (CTP). Notice Announcing Availability of Common Data Elements for Tobacco Regulatory Research via the PhenX Toolkit, 2017. Available: https://grants.nih.gov/grants/guide/notice-files/NOT-OD-17-034.html [Accessed 9 Apr 2019].
9 Piper ME, Brown DC, Hendershot TP, *et al*. PhenX: host: Social/cognitive measures for tobacco regulatory research. *Tob Control* 2019;28:s143–50.
10 Giovino GA, Swan GE, Blount BC, *et al*. PhenX: host: Biobehavioral measures for tobacco regulatory research. *Tob Control* 2019;28:s151–7.
11 O'Connor RJ, Watson CH, Swan GE, *et al*. PhenX: agent measures for tobacco regulatory research. *Tob Control* 2019;28:s158–64.
12 Ribisl KM, Chaloupka FJ, Kirchner T, *et al*. PhenX: new vector measures for tobacco regulatory research. *Tob Control* 2019;28:s165–72.
13 Unger JB, Chaloupka FJ, Vallone D, *et al*. PhenX: environment measures for tobacco regulatory research. *Tob Control* 2019;28:s173–80.
14 Garcia-Cazarin M, Mandal R, Mayne RG, *et al*. Host-agent-vector- environment measures in electronic cigarette research. *Tob Control* 2019;28:s181–7.
15 Kaufman AR, Persoskie A, Twesten J, *et al*. A review of risk perception measurement in tobacco control research. *Tob Control* 2019;28:s188–96.
16 Conway KP, Vullo GC, Kennedy AP, *et al*. Data compatibility in the addiction sciences: an examination of measure commonality. *Drug Alcohol Depend* 2014;141:153–8.