Rapid Collection of Experimental Physicochemical Property Data to Inform Various Models and Testing Methods

American Chemical Society
255th National Meeting
New Orleans, LA

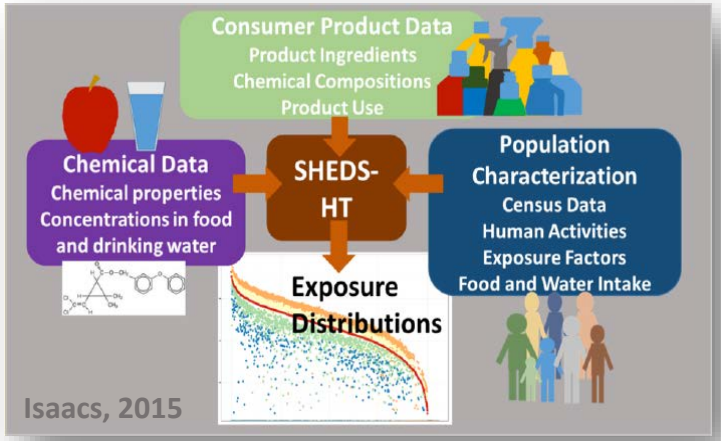
March 21, 2017

Chantel I. Nicolas, PhD
Associate Investigator
919-558-1292
cnicolas@scitovation.com
www.scitovation.com

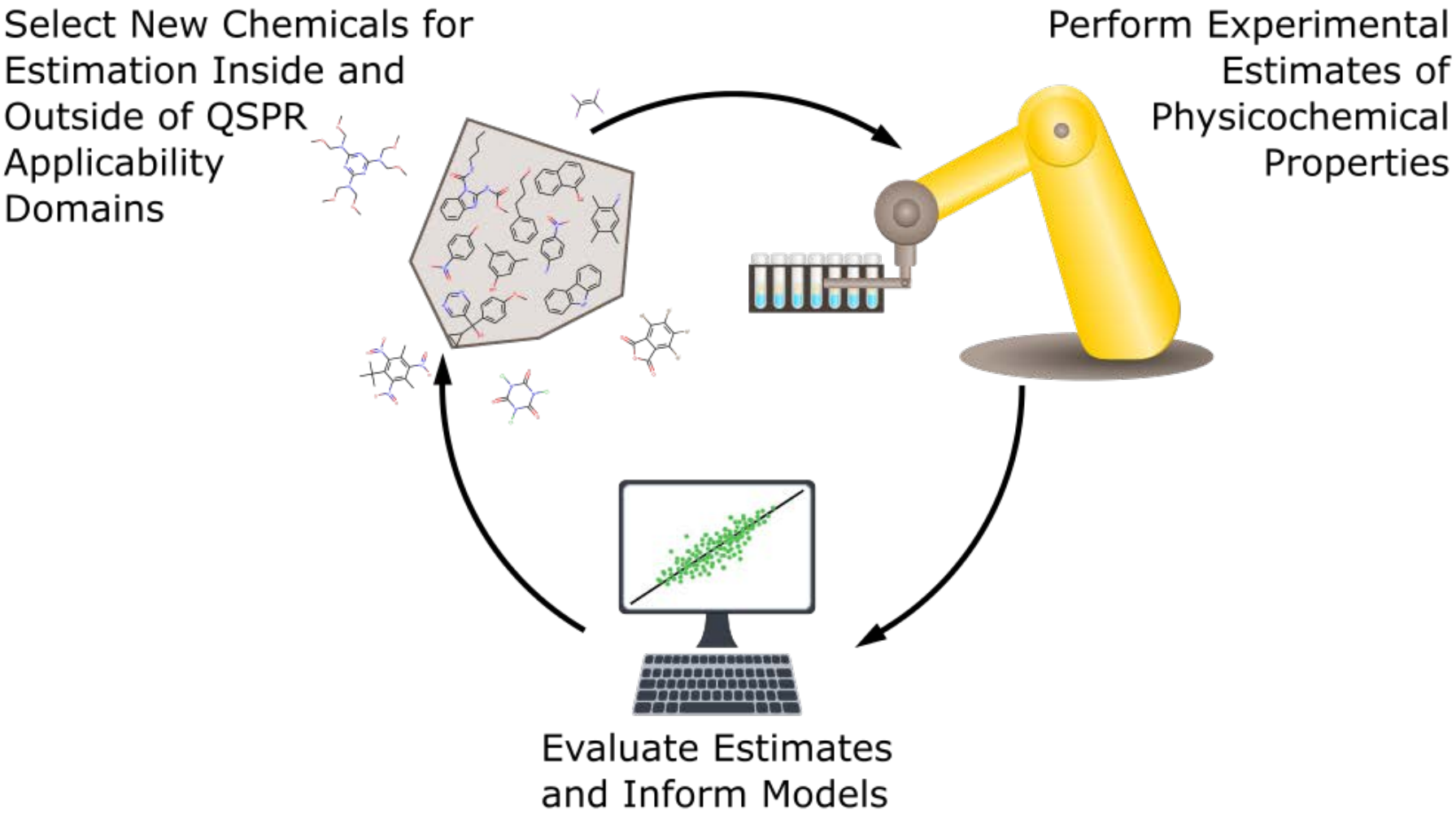
ScitoVation
INNOVATIVE CELL BASED SCIENCE

The Need for Reliable Physicochemical Properties

- High-throughput models have been developed to estimate exposure and dose of substances from near-field sources
- Physicochemical properties parameterize many of these models
- Quantitative structure property relationship (QSPR) models to predict these properties have been developed to fill data gaps
- Most measured datasets of "biologically relevant" physicochemical properties are for substances relevant to the pharmaceutical industry



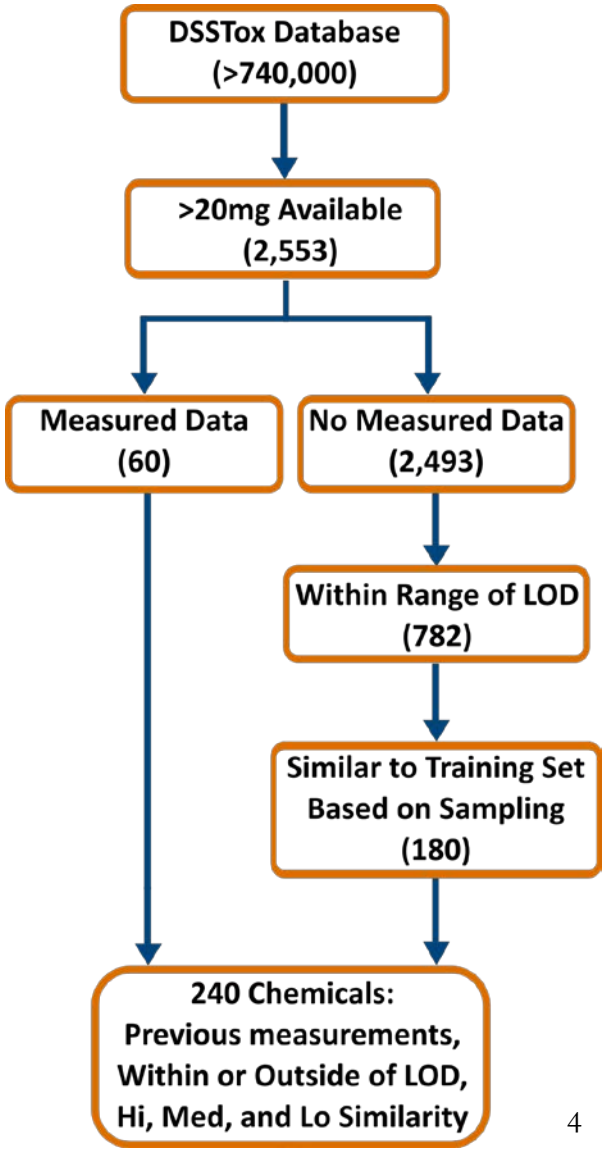
Continuous Updating of QSPR Models



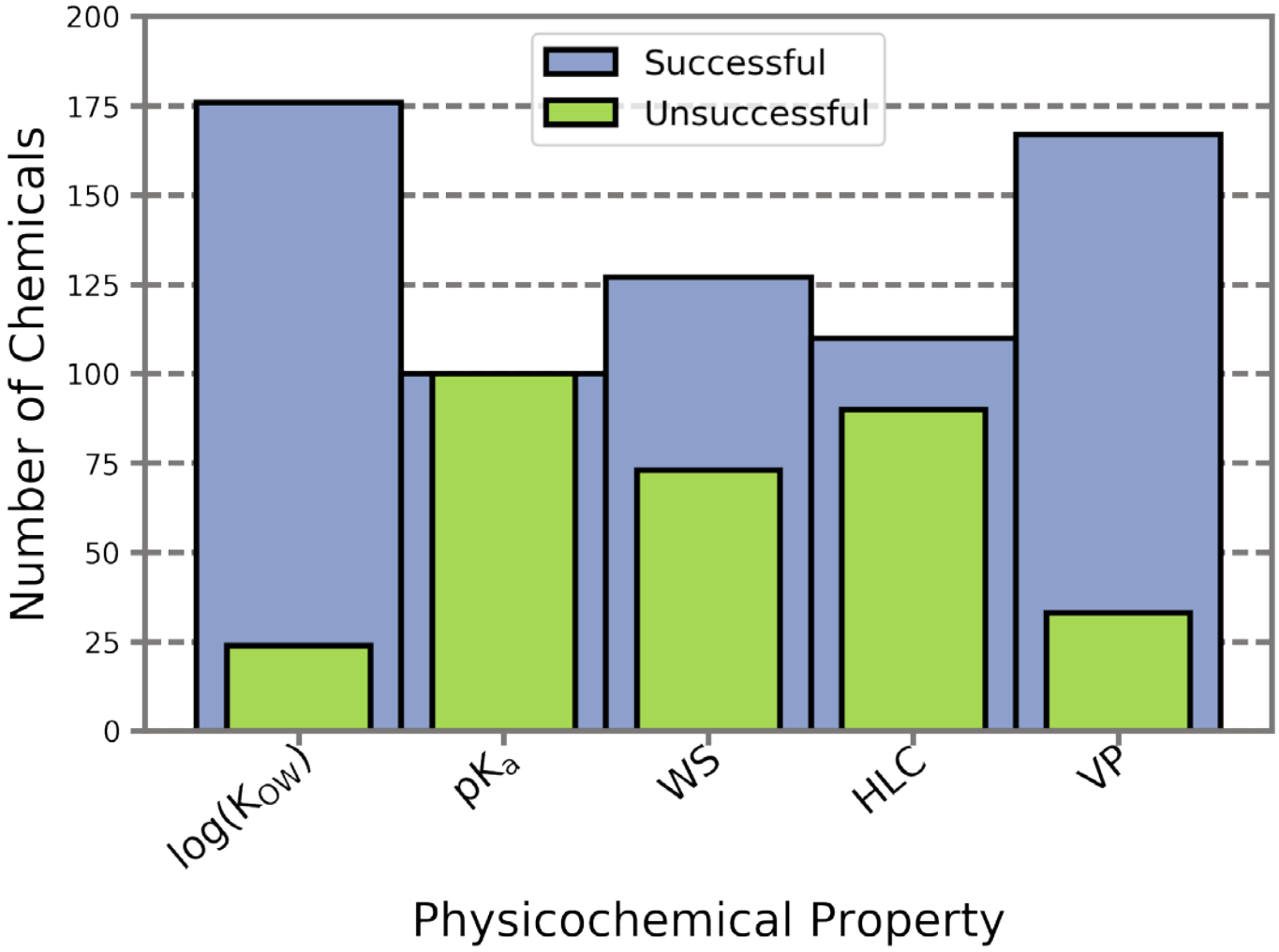
- New measured data are needed for improving QSAR applicability domains

Process for Selecting Pilot Chemicals for Testing

- If we select the same chemicals in current training set (PHYSPROP), then we get a nice comparison, but we get no expansion of chemical space
- If we select all new chemicals, then we get no comparison, but an expansion of chemical space
- For the pilot, 200 chemicals were submitted for measurement of 5 properties:
  - $Log_{10}K_{ow}$
  - Henry's Law constant (HLC)
  - Vapor pressure (VP)
  - Water solubility (WS)
  - Acid dissociation constant (pK_a)

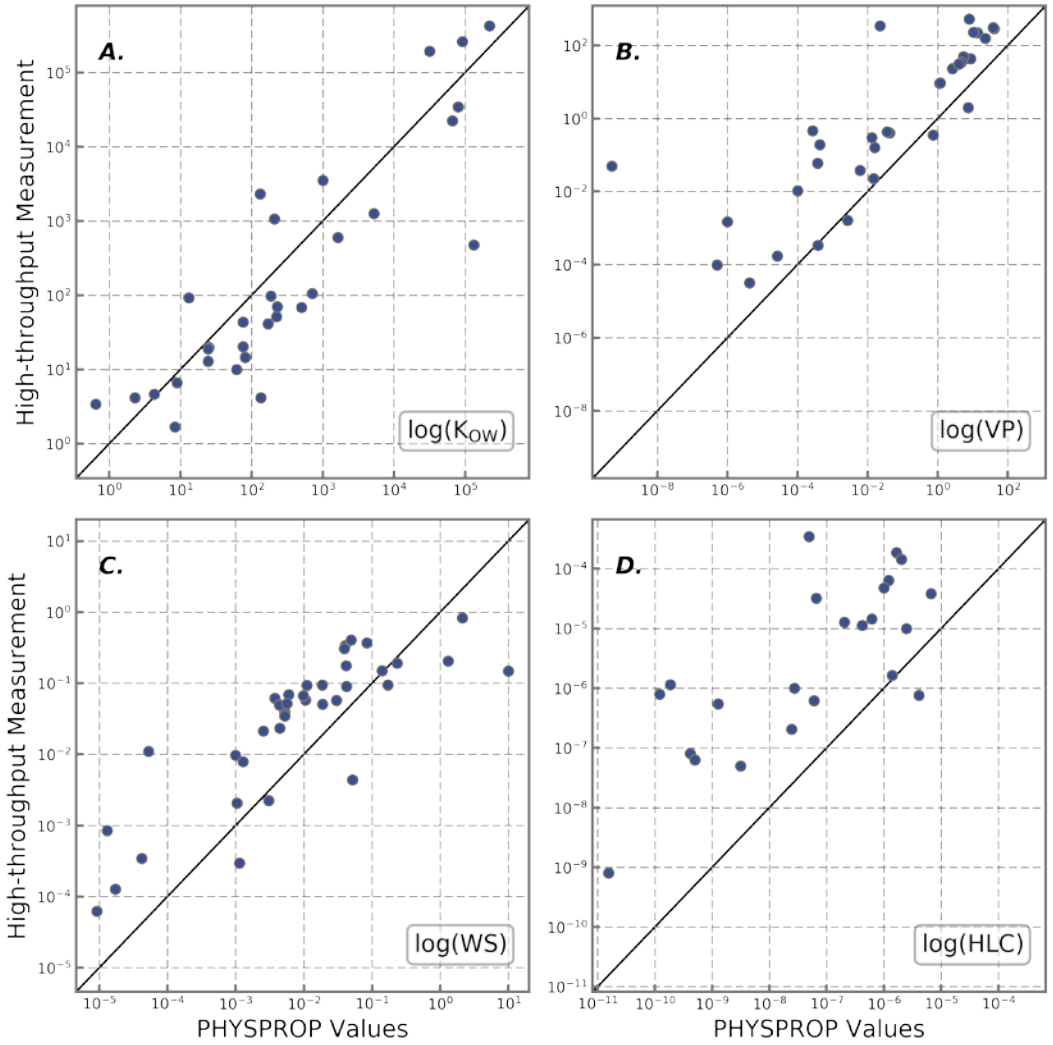


Summary of Successful Measurements



- $Log_{10}K_{ow}$ returned the greatest number of measurements and WS, the least

Success for Compounds in Existing Training Set



- Correspondence between new measurements and those previously reported

QSPR Tools and the Property Predictions Used

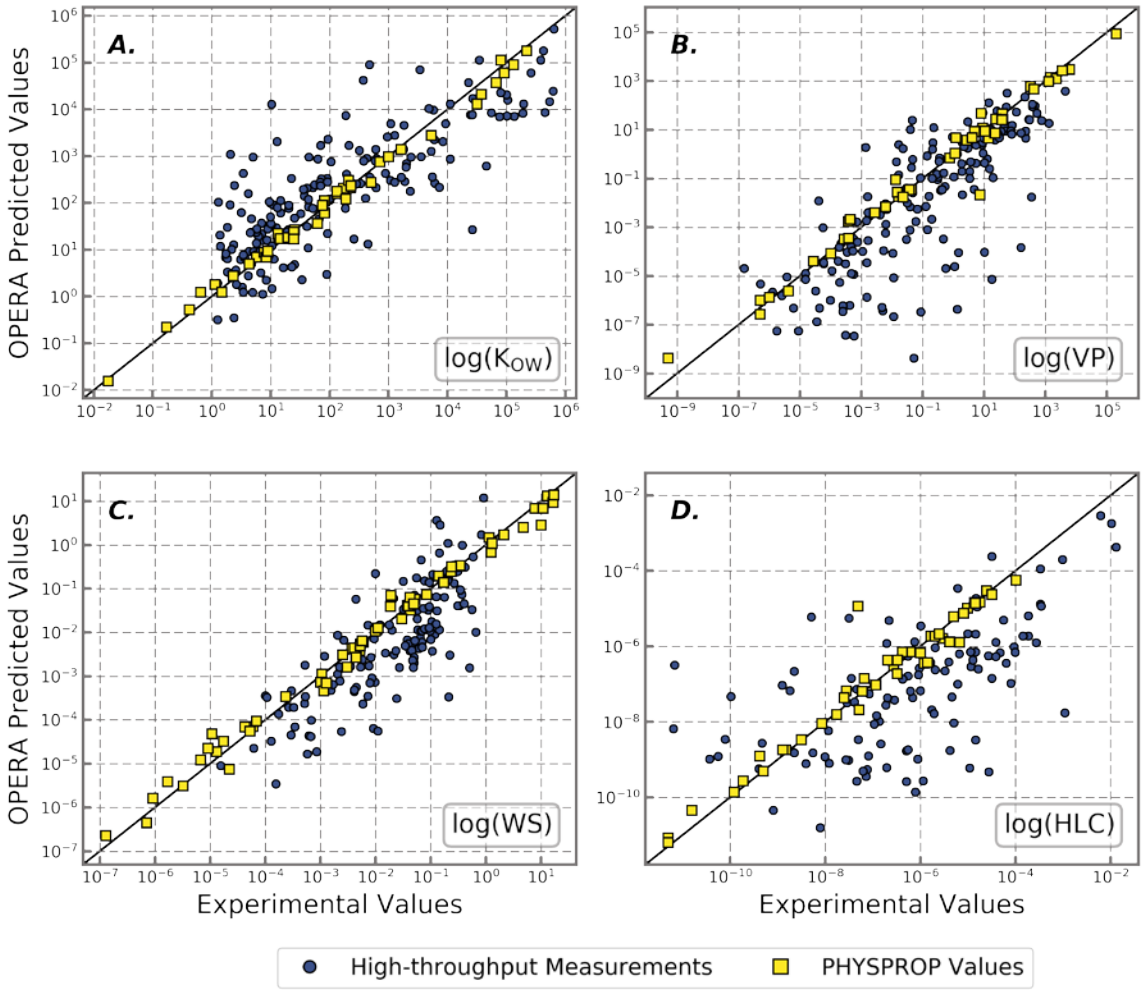
- ACD/Labs
  - Advanced Chemistry Development, Inc.
  - $Log_{10}K_{ow}$, WS, and pK_a
- ChemProp
  - Chemical Properties Estimation Software System (UFZ)
  - $Log_{10}K_{ow}$, VP, and WS
- EPISuite
  - Estimation Program Interface
  - $Log_{10}K_{ow}$, VP, WS, and HLC
- NICEATM
  - NTP Interagency Center for the Evaluation of Alternative Toxicological Methods
  - $Log_{10}K_{ow}$, and WS
- OPERA
  - OPEn (quantitative) Structure-activity Relationship Application
  - $Log_{10}K_{ow}$, VP, WS, and HLC

Trends Across Modeling Tools

		RMSE					
	N	ACD/Labs	ChemProp*	EPI Suite	NICEATM	OPERA	PHYSPROP
$LogK_{ow}$	32	8.30E-01	7.91E-01	8.02E-01	7.75E-01	7.36E-01	7.71E-01
VP	33	-	1.44E+02	1.38E+02	-	1.36E+02	1.38E+02
HLC	23	-	-	5.99E-04	-	8.77E-05	8.96E-05
WS	36	1.20E+00	2.33E-01	3.07E-01	2.28E-01	5.07E-01	1.67E+00
pKa†	76	6.15E+00	-	-	-	-	-

- Root mean squared errors (RMSE) are reported to emphasize average deviations between new values and previous or predicted values.
- The trends are as follows:
  - $Log_{10}K_{ow}$: OPERA < NICEATM < ChemProp < EPI Suite < ACD/Labs
  - VP: OPERA < EPI Suite < ChemProp
  - WS: ACD/Labs < NICEATM < ChemProp < EPI Suite < OPERA
  - HLC: OPERA < EPI Suite

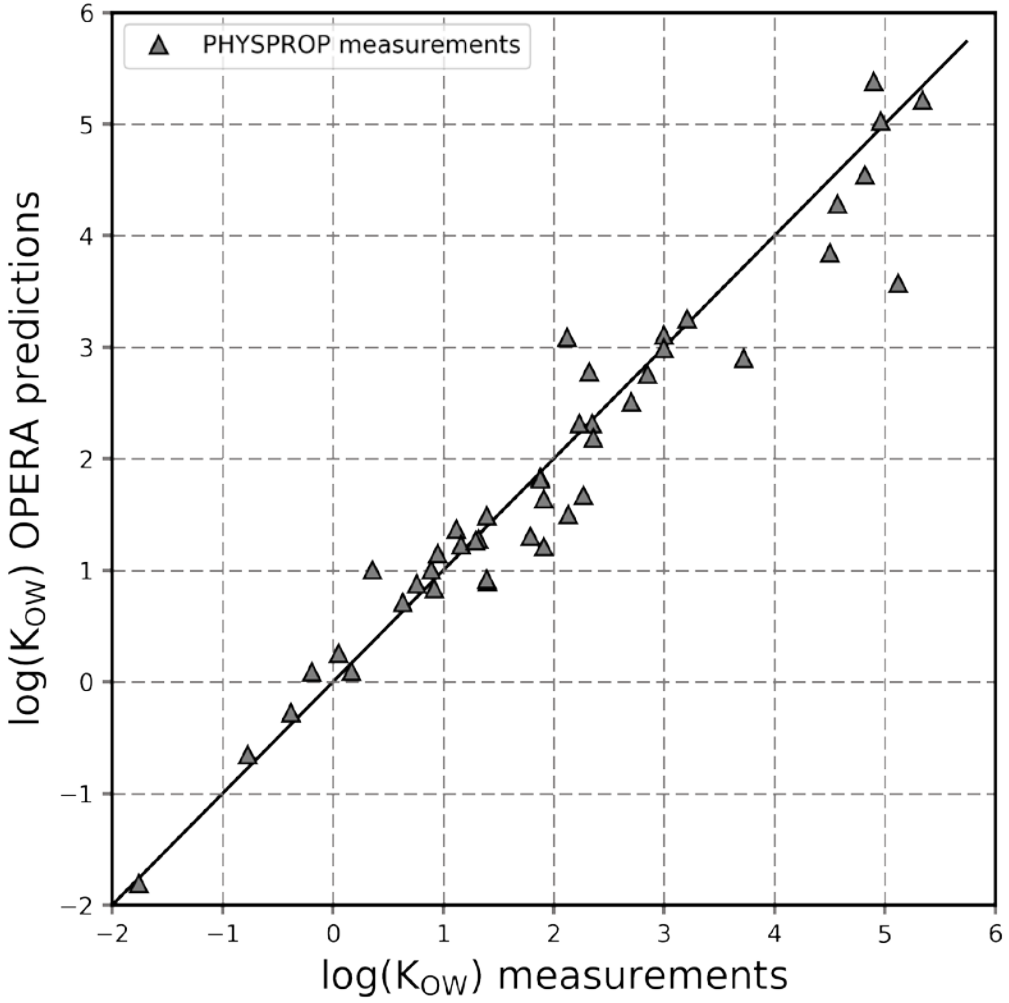
OPERA Trends



- Comparison of OPERA predictions to both new and previous measurements

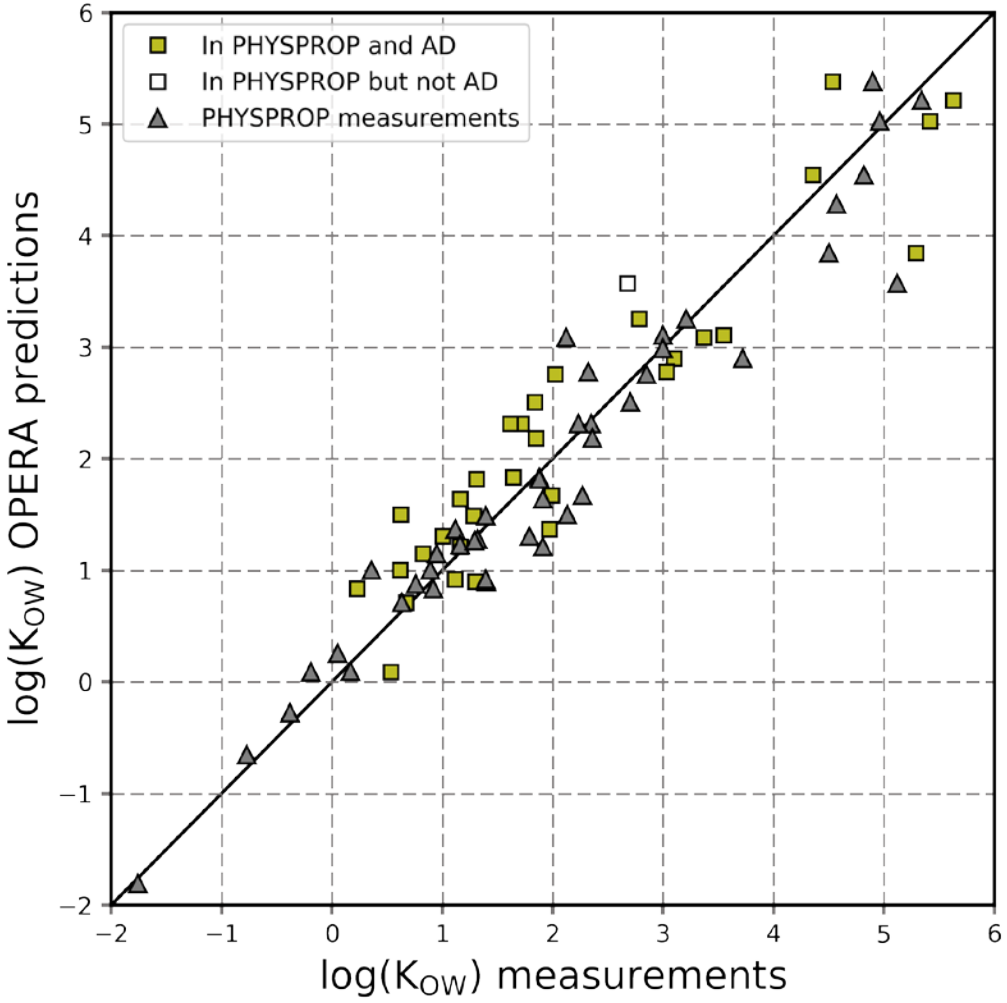
OPERA $Log_{10}K_{ow}$ Trends and Applicability Domain

- Chemicals not in PHYSPROP were not in OPERA training set
- Chemicals with lower $log_{10}K_{ow}$ values were unamenable to new measurements



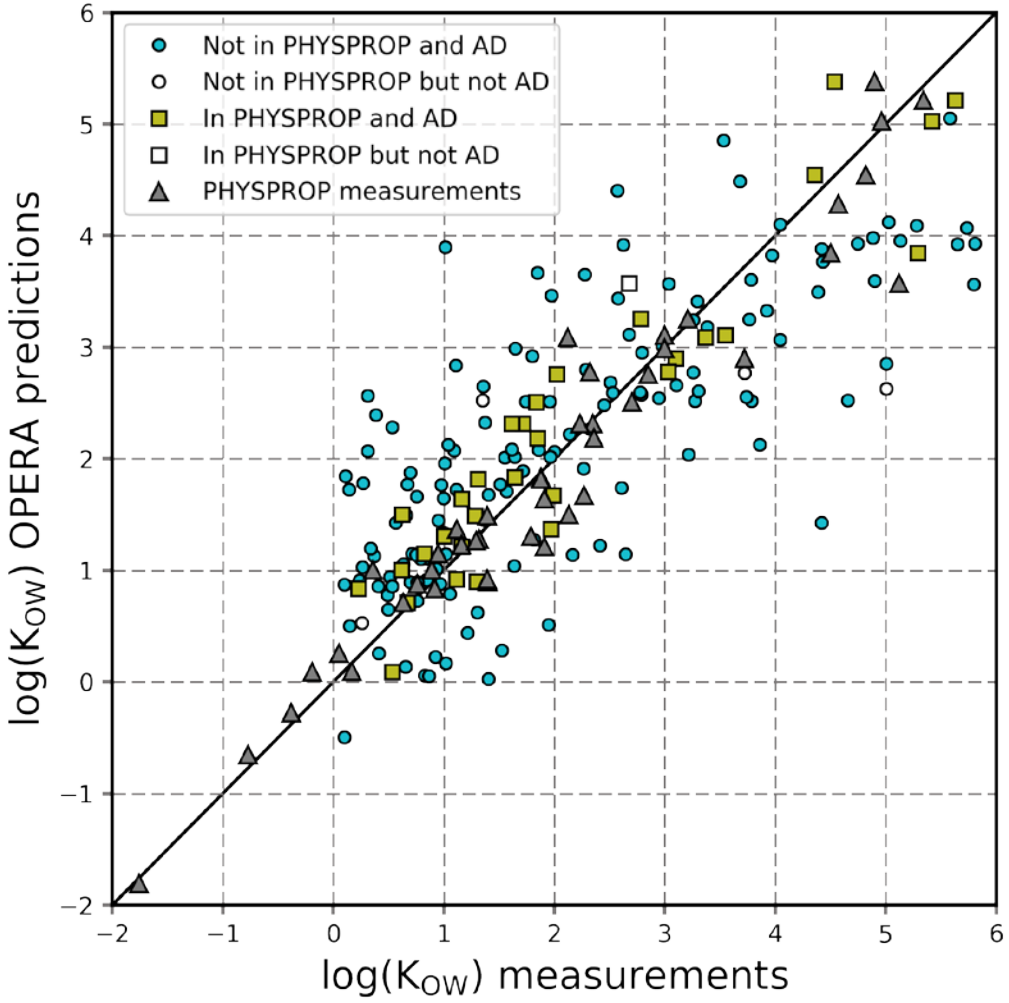
OPERA $Log_{10}K_{ow}$ Trends and Applicability Domain

- Chemicals not in PHYSPROP were not in OPERA training set
- Chemicals with lower $log_{10}K_{ow}$ values were unamenable to new measurements

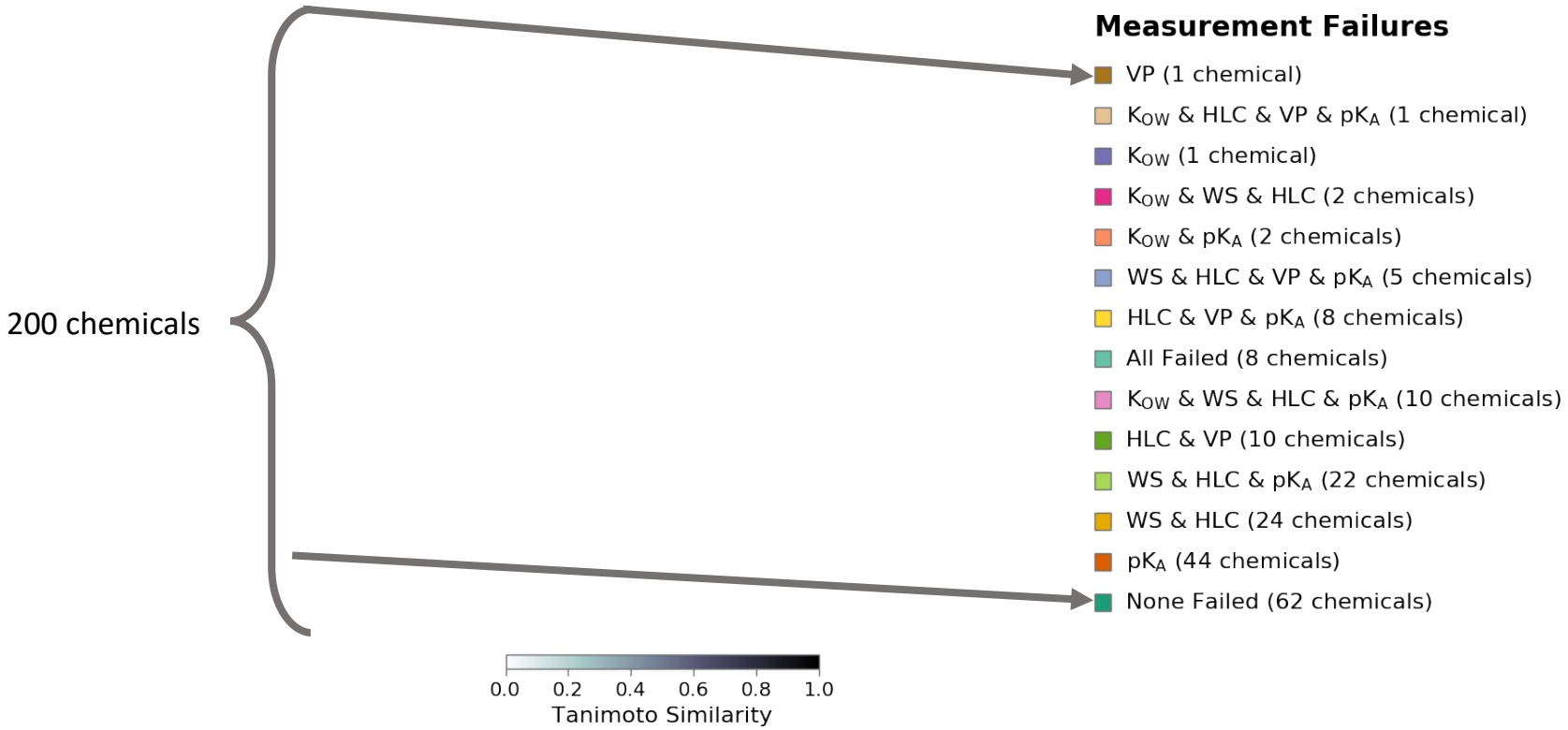


OPERA $Log_{10}K_{ow}$ Trends and Applicability Domain

- Chemicals not in PHYSPROP were not in OPERA training set
- Chemicals with lower $log_{10}K_{ow}$ values were unamenable to new measurements

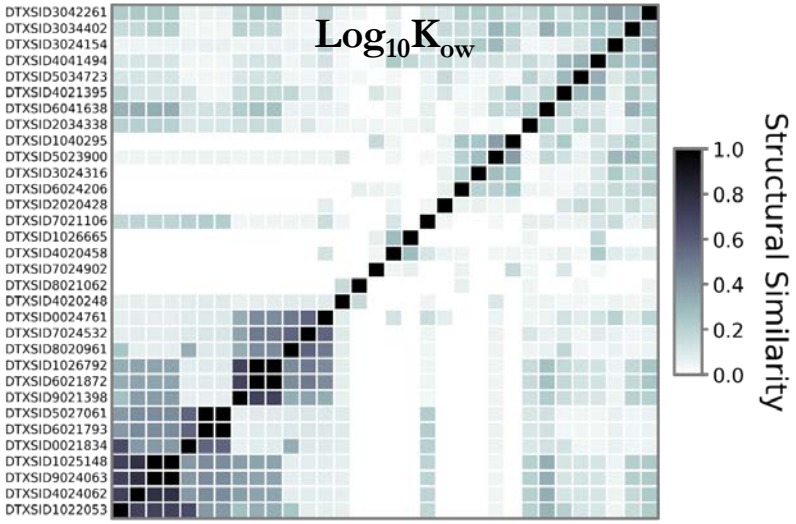


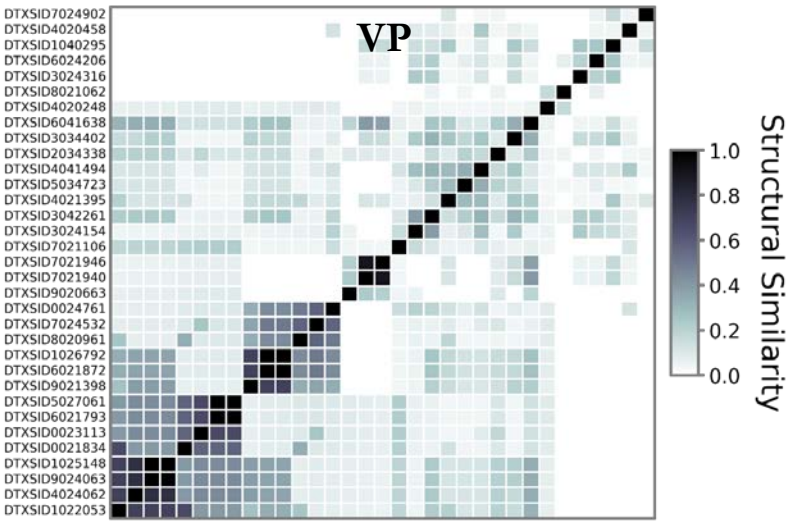
Structure Similarity for Failures based by Property

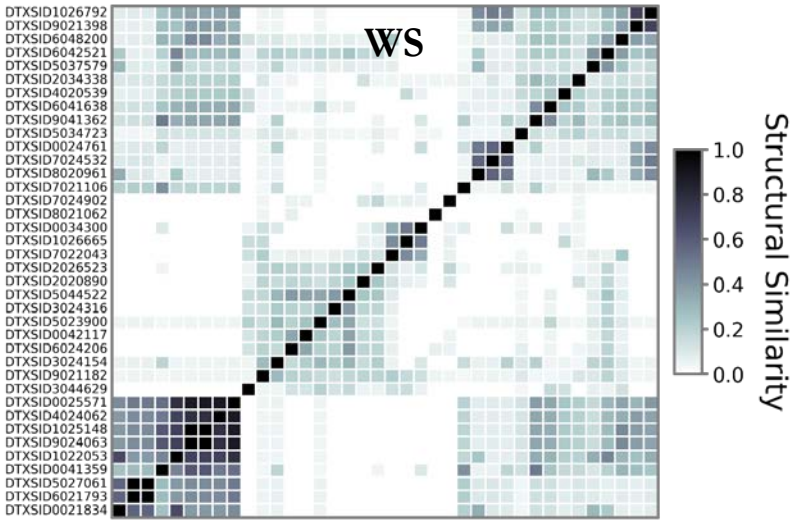


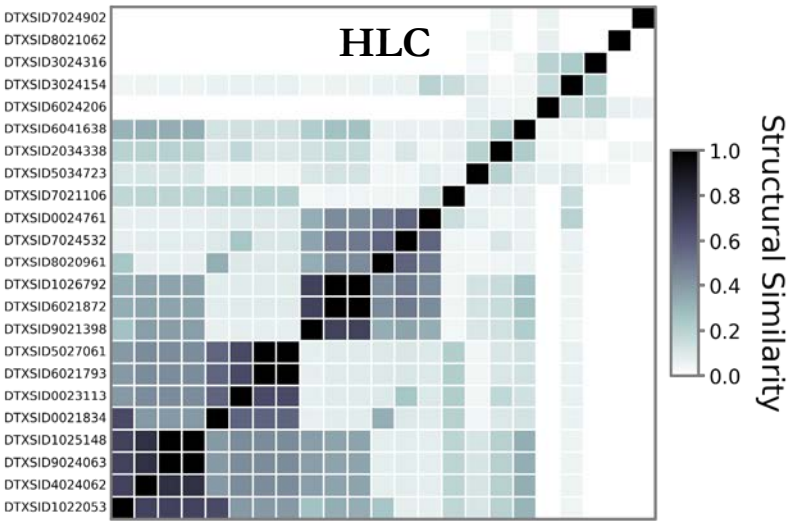
- Chemotypes were used to seek patterns in measurement failures across properties

Structural Similarity For New vs Previous Values



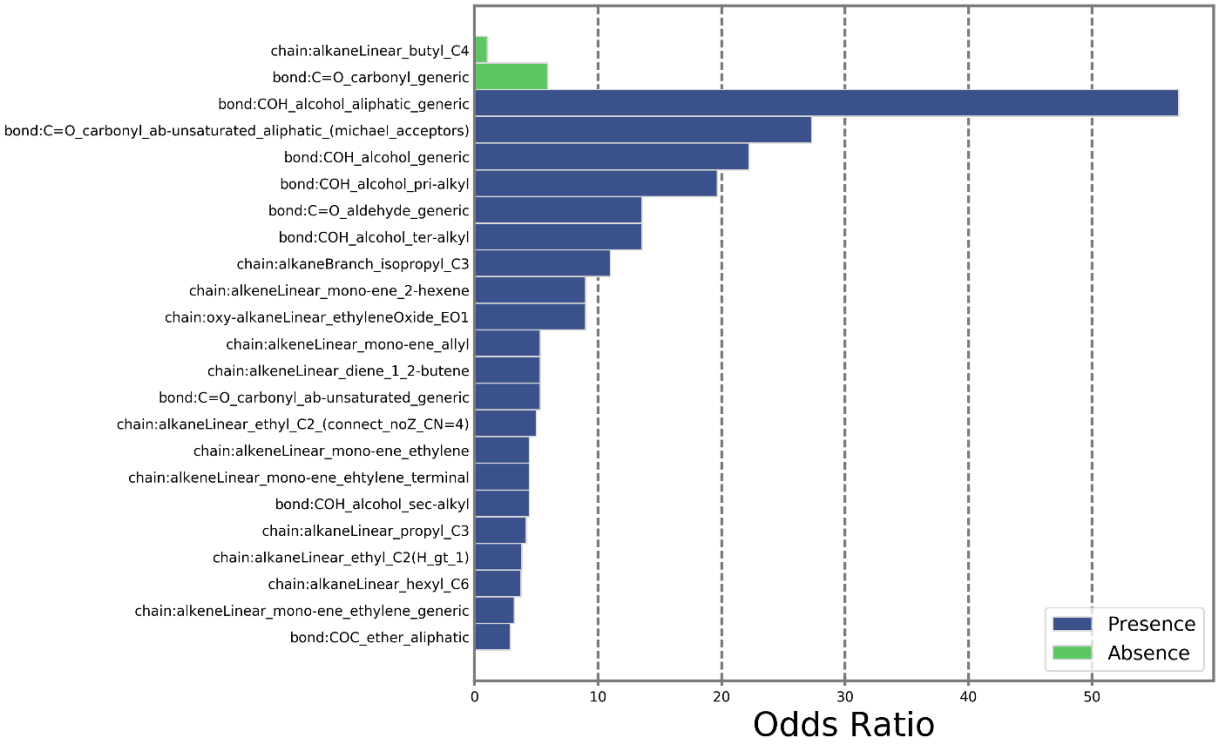






- Structural similarity between chemicals from PHYSPROP and those with new measurements

Odds Ratios for Successful Measurements



- Odds Ratio give the odds that a substructure contributes to failure of all measurements
  - Odds Ratio > 1 → positive association
  - Odds Ratio = 1 → no association
  - Odds Ratio < 1 → negative association
- There were 21 substructures with positive associations
- There were 2 substructures with negative associations

Conclusion

- Five physicochemical properties have been measured in a high-throughput fashion
- A selection process was implemented that allowed for diverse selection of 200 compounds for measurement
- Results from the new measurements were compared against both previously measured values and a collection of QSAR models that predict these values
- These new measurements can help improve chemical space covered by QSAR thus improving parameters for exposure and dose models

Acknowledgements

ScitoVation

Melvin Andersen
Tino Balbuena
Tyler Beames
Michael Black
Harvey Clewell
Rebecca Clewell
Michael Dzierlenga
Saad Haider
Eric Hack
Jessica Hartman
Kamel Mansouri
Patrick McMullen
Marjorie Moreau
Melyssa Minto
Chantel Nicolas
Salil Pendse
Martin Phillips
Miyoung Yoon

Collaborators

USEPA
Cody Addington
Peter Egeghy
Chris Grulke
Kristin Isaacs
Jeremy Leonard
Grace Patlewicz
Robert Pearce
Katherine Phillips
Jim Rabinowitz (Retired)
Ann Richard
Cecilia Tan
Russell Thomas
Dan Stout
Cory Strope
John Wambaugh
Barbara Wetmore
Antony Williams

Cosmetics Europe
Bertrand Desprez
Sabrina Dourte
Martina Klaric
Nicola Hewitt

Proctor & Gamble
Corie Ellison
Sarah Tozer

Ramboll-Environ
Harvey Clewell
Jerry Campbell

Southwest Research Institute
Alice Yau

ToxStrategies
Rebecca Clewell
Miyoung Yoon

UNC – Chapel Hill
Alexander Tropsha

Funding Sources

American Chemistry Council
Cosmetics Europe

This work does not necessarily reflect U.S. EPA policy.