EPA United States Environmental Protection Agency

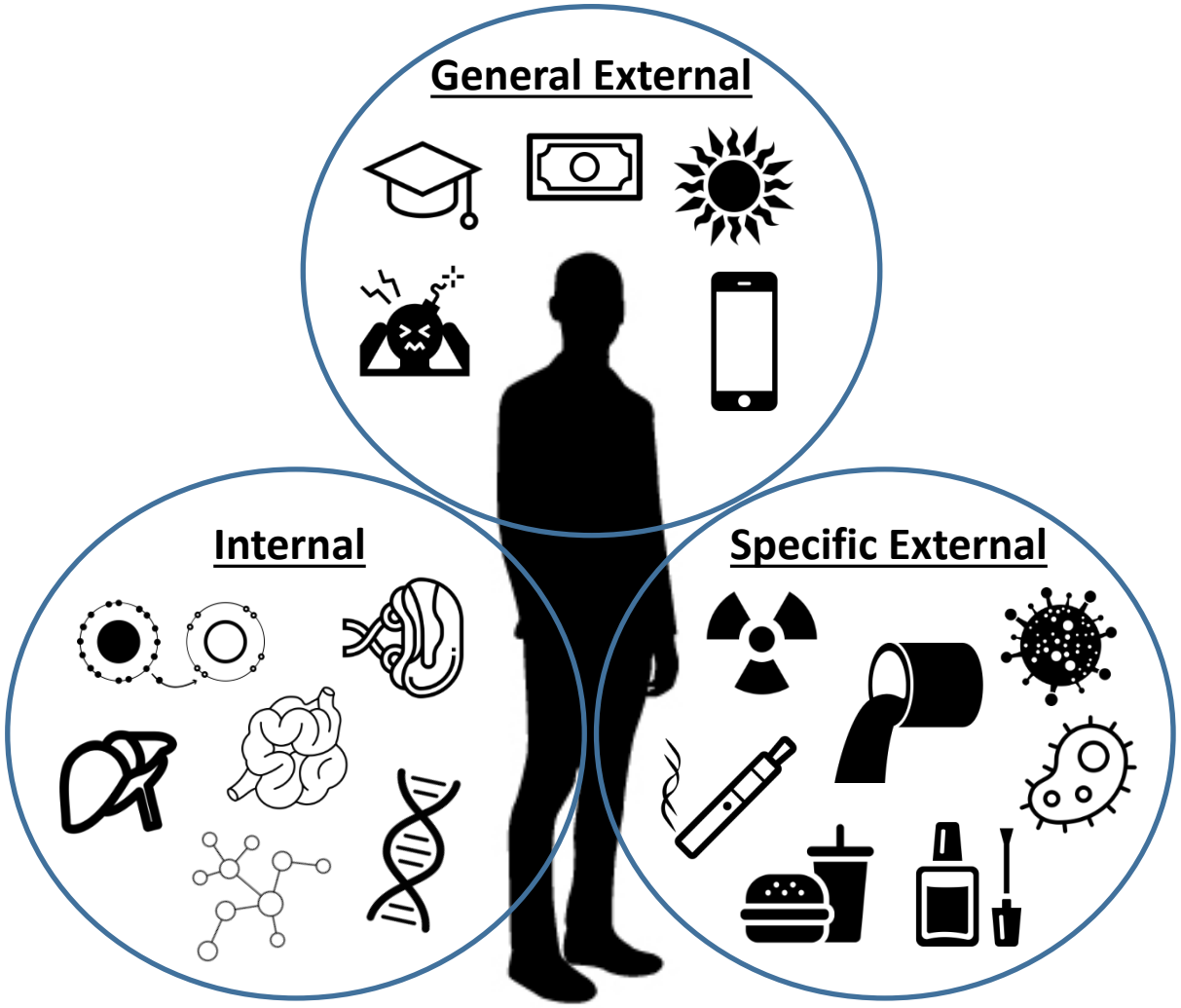
Biomonitoring in the 21st Century: Challenges and Opportunities

Jon R. Sobus, Ph.D.

Office of Research and Development
Center for Computational Toxicology and Exposure

July 13, 2021

“Exposure” is Extraordinarily Complex!

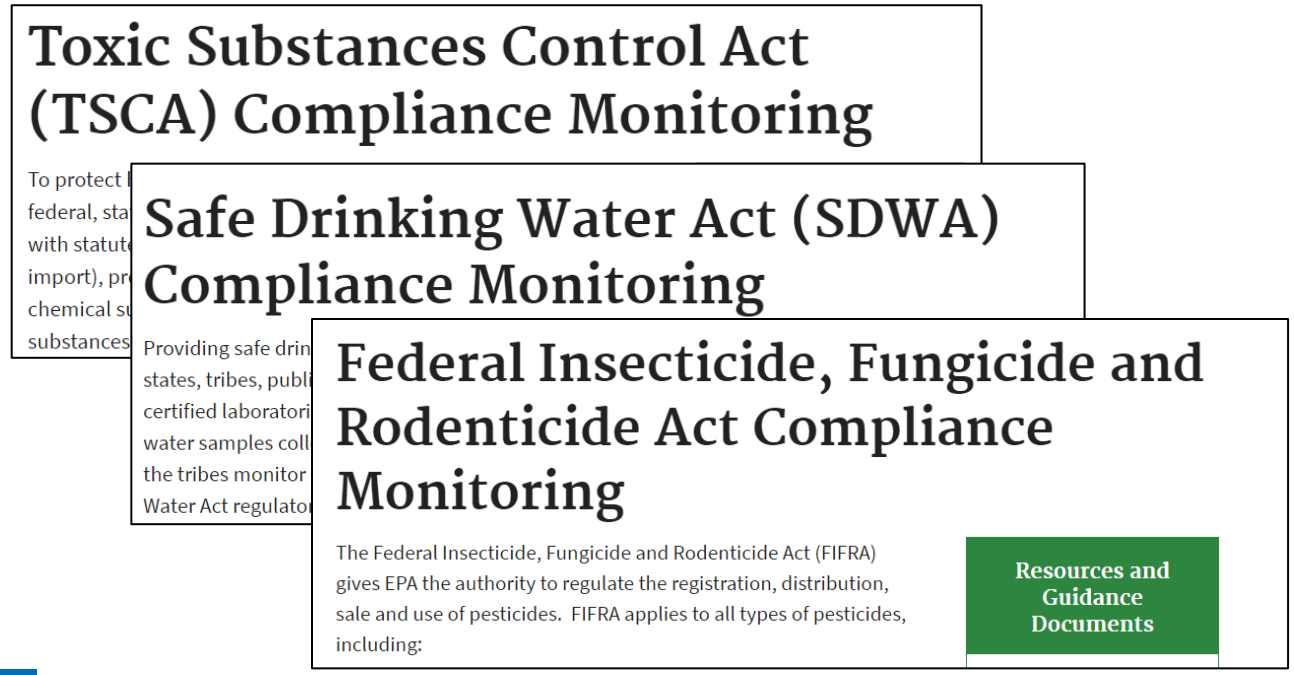


- **Stressors:**
  - External: Chemical, biological, radiological, diet, stress
    - Complex mixtures
  - Internal: Endocrine activity, metabolic activity
- **Sources:**
  - Outdoor air, indoor (residential & workplace) air
  - Food, beverages, food packaging, dust, soil
  - Consumer products (lotions, sprays, cleaning products)
- **Pathways & Routes:**
  - Inhalation, ingestion (dietary & non-dietary), dermal
  - Aggregate: one stressor, multiple routes
  - Cumulative: multiple stressors, multiple routes
- **Frequency × Duration × Intensity**
- Individual receptors can affect exposure
  - External → **vulnerable** populations (economic & age)
- Individual receptors can react differently to exposure
  - Internal → inherent/genetic/epigenic **susceptibility**

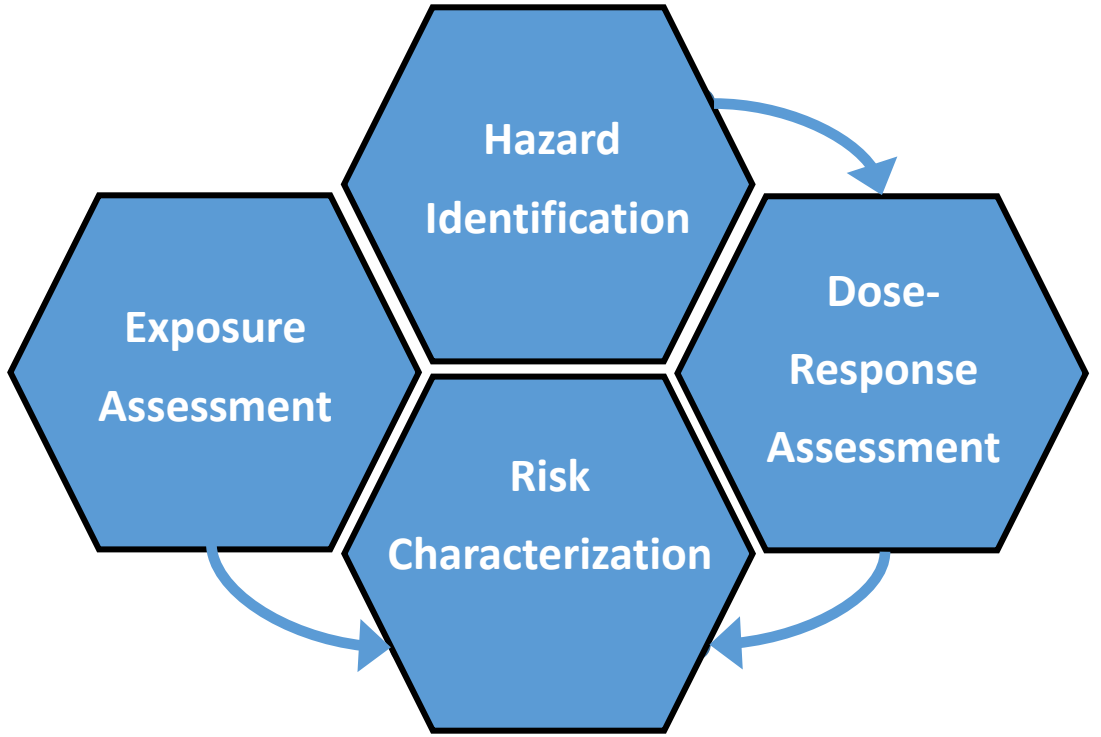
Figure adapted from: Wild CP. *Int. J. Epidemiol.* 2012 Feb;41(1):24-32

Measurement Data Used to Monitor Stressors

- **Measurement data needed to:**
  - Characterize risk from chemical stressors
  - Regulate chemical use & disposal
  - Manage human & ecological exposures
  - Ensure compliance under federal statutes



Components of the Risk Assessment Process



Comparing Monitoring Strategies

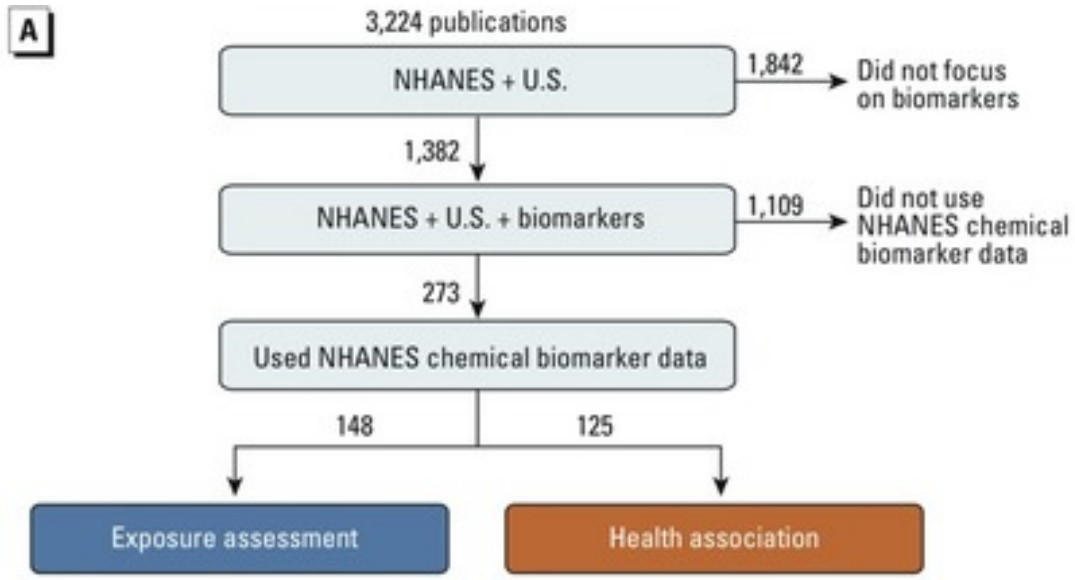
- Bottom-Up → Measure important exposures outside of the receptor(s)
  - Media: drinking water, food, air, soil, consumer products, etc.
  - Advantage: directly link the stressor(s) to the source(s); best for mitigation
  - Disadvantage: MANY possible stressors/media combinations

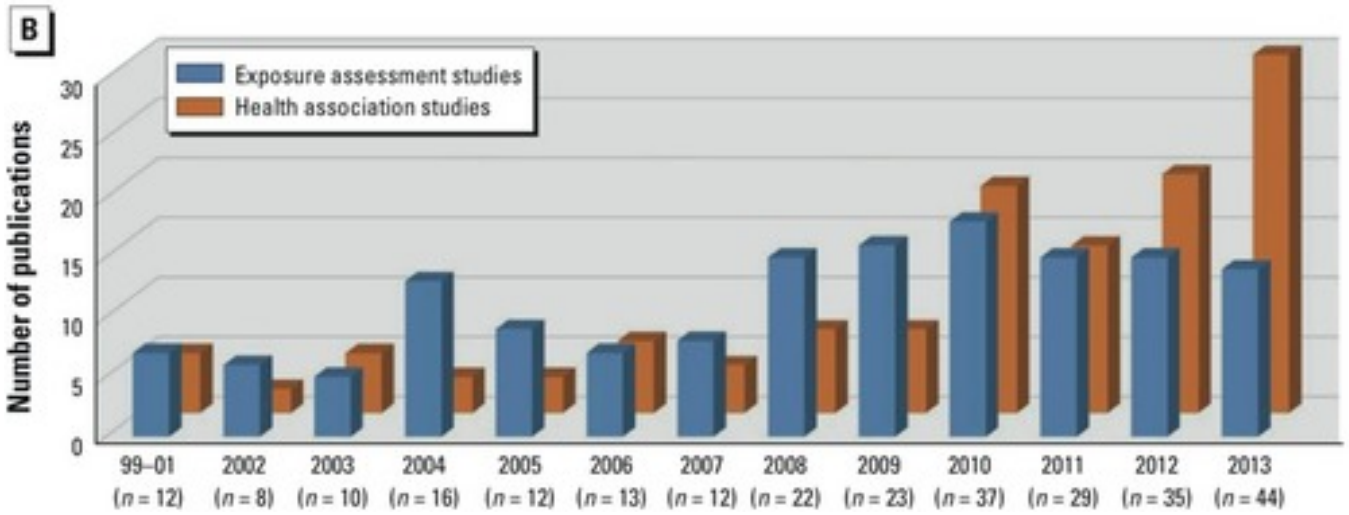
- Top-Down → Measure important exposures within the receptor(s)
  - Media: blood, urine, breath, hair, nails, teeth, target tissue, etc.
  - Advantage:
    - (1) Biomarkers can integrate exposure overall "all" pathways & routes
    - (2) Closer to "target" → simultaneous evaluation of exposure and biological response
  - Disadvantages
    - (1) Biomarker levels affected by exposure-independent processes (e.g., hydration, lipid levels)
    - (2) Requires linkage to source for regulatory consideration & mitigation
    - (3) Invasive sampling

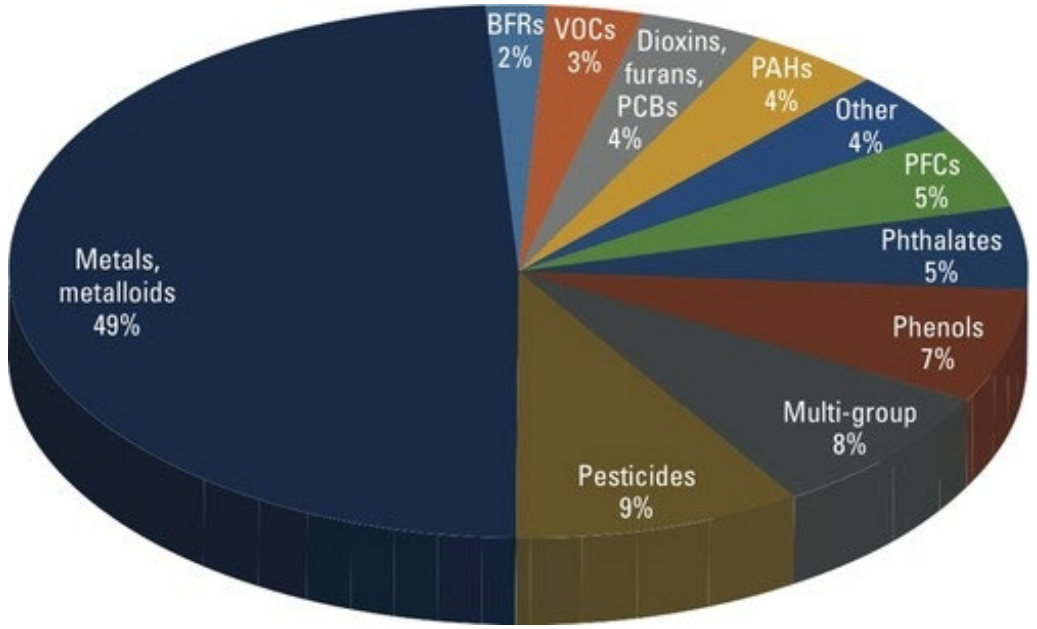
Targeted Biomonitoring via CDC's NHANES

- Representative sample of U.S. (civilian, non-institutionalized) pop.
- Thousands of participants per year
- Running from 1971 – present
- ~30 chemical biomarkers (1999) to ~400 chemical biomarkers (2021)
- "Spot sampling" of blood and urine
- Exposure-related metadata
- Health-related metadata and biological measures
- Intended use → time-based exposure trends
- Limitations:
  - Not all biomarkers measured in all participants
  - Limited access to geographical information
  - Cross-sectional design

Actual Uses of NHANES Biomarker Data







A Section 508–conformant HTML version of this article is available at http://dx.doi.org/10.1289/ehp.1409177.

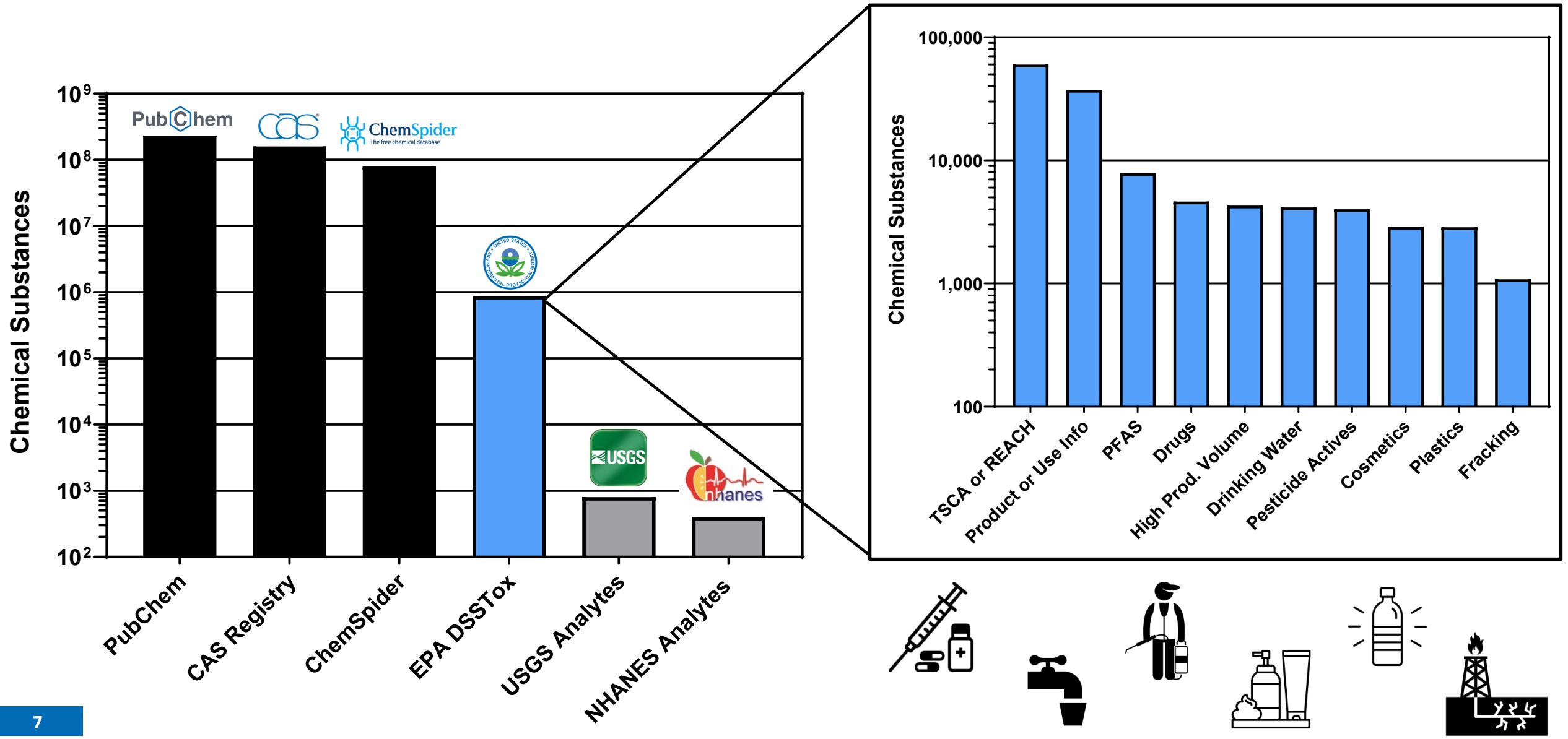
Review

Uses of NHANES Biomarker Data for Chemical Risk Assessment: Trends, Challenges, and Opportunities

Jon R. Sobus,[1] Robert S. DeWoskin,[2] Yu-Mei Tan,[1] Joachim D. Pleil,[1] Martin Blake Phillips,[3] Barbara Jane George,[4] Krista Christensen,[5] Dina M. Schreinemachers,[6] Marc A. Williams,[4] Elaine A. Cohen Hubal,[7] and Stephen W. Edwards[4]

[1]National Exposure Research Laboratory, and [2]National Center for Environmental Assessment, U.S. Environmental Protection Agency (EPA), Research Triangle Park, North Carolina, USA; [3]Oak Ridge Institute for Science and Education (ORISE) Participant, Research Triangle Park, North Carolina, USA; [4]National Health and Environmental Effects Research Laboratory, U.S. EPA, Research Triangle Park, North Carolina, USA; [5]National Center for Environmental Assessment, U.S. EPA, Washington, DC, USA; [6]National Health and Environmental Effects Research Laboratory, U.S. EPA, Chapel Hill, North Carolina, USA; [7]Office of Research and Development, U.S. EPA, Research Triangle Park, North Carolina, USA

Data Disparity: Have vs. Need



Challenges with Targeted Monitoring

- High-quality monitoring data are unavailable for most chemicals
- Measurement data traditionally generated using “targeted” methods
- Targeted analytical methods:
  - Require *a priori* knowledge of chemicals of interest
  - Produce data for few selected analytes (generally 10s)
  - Require standards for method development & compound quantitation
  - Are blind to emerging contaminants
  - Can’t keep pace with the needs of 21st century chemical safety evaluations

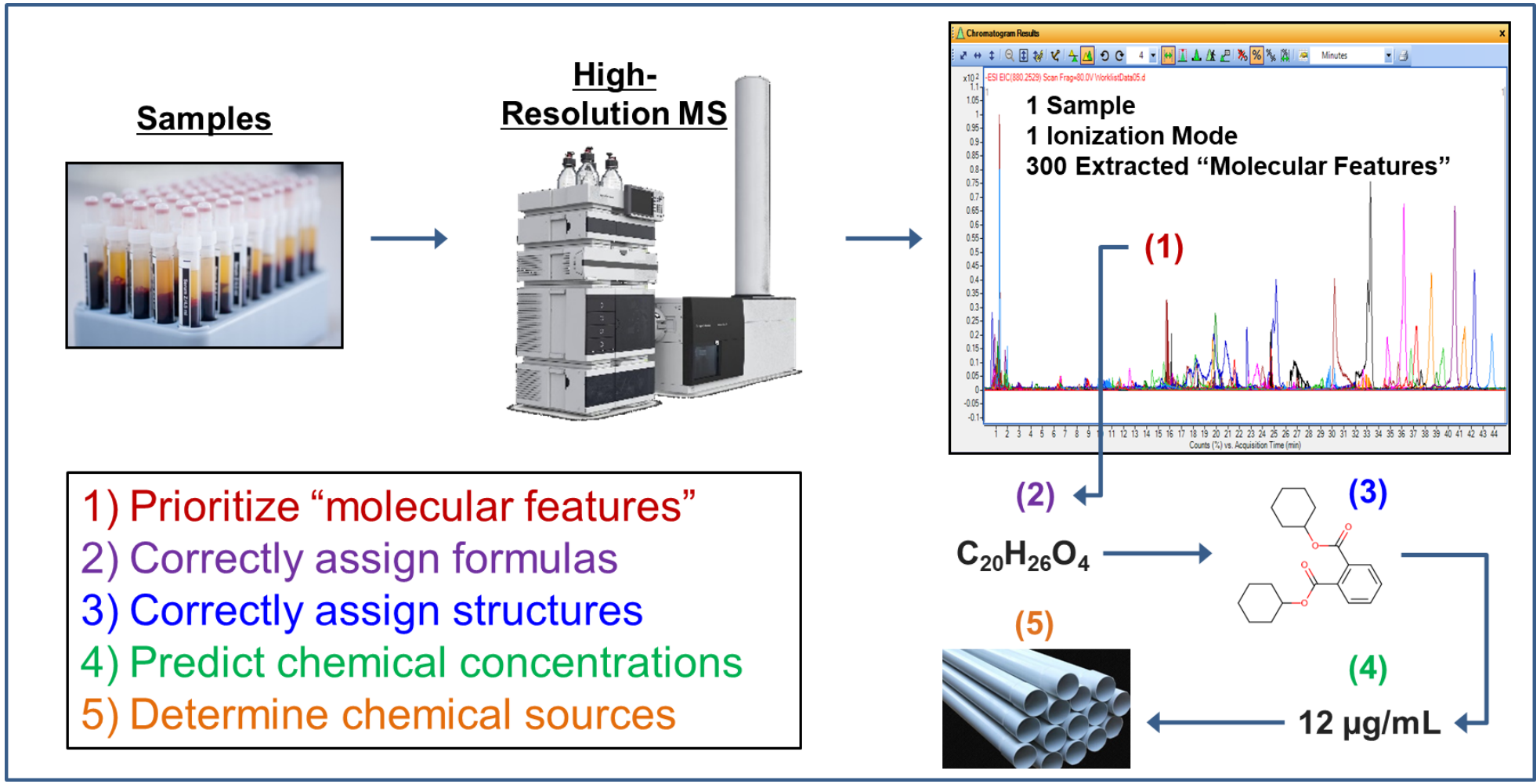
Shrinking Data Gaps with Non-Targeted Analysis

Rapidly screen for "knowns"

Discover "unknowns"

Uncover historical exposures

Link stressors to biological responses



Developing and Disseminating Guidance Materials

- BP4NTA → Borne out of 2018 EPA NTA workshop
- ~100 U.S. and international members
  - Government, academia, and industry
- Working Group Objectives:
  - Short term → define common NTA terms, concepts, and performance metrics
  - Short term → provide recommendations on research & reporting best practices
  - Long term → enable proficiency testing
- Products (including 3 manuscripts):
  - Website with key resources and links: https://nontargetedanalysis.org/
  - Guidance documents with definitions & supporting info
  - "NTA Study Reporting Tool" to standardize reporting (proposals & manuscripts)

The NTA Study Reporting Tool

NTA Study Chronology

Study Sections & Categories			Example Information to Report	Numeric & Colorimetric Scoring	Rationale/Notes
Methods	Study Design	Objectives & Scope	• Study goals and hypotheses • Scope of the study with respect to use of NTA / suspect screening • Expected chemical coverage of approach and potential limitations	1	
		Sample Information & Preparation	• Sample collection/replication, handling/storage, preparation, extraction, & clean-up methods (and related QA practices) • Intended use of samples (e.g., method development, compound identification, etc.) • Development and intended use of blanks	2	
		QC Spikes & Controls	• Development of spikes/controls (e.g., isotopically labeled standards/spikes, native standard spikes, matrix pools) • Intended use of QC or other spikes/controls (e.g., to monitor instrument performance, data normalization, etc.)	2	
	Data Acquisition	Analytical Sequence	• Sample randomization and use of replicate injections • Inclusion of blanks and QC samples in the acquisition sequence • Information about single vs. multiple analytical batches	3	
		Chromatography	• Instrument specifications • Method settings (e.g., column/guard, mobile phases, gradient, injection techniques)	3	
		Mass Spectrometry	• Instrument specifications • Instrument calibration and/or tuning procedures • Method settings (e.g.,	3	
			• File conversion inform • Software program(s) u • Workflow steps (e.g., p • Feature detection thres • Data correction or nor	2	
			• Software programs(s) p • Basic statistical analys settings/thresholds • Chemometric analysis	NA	
			• Software program(s) u • Libraries and databases • Workflow steps (e.g., p • Workflow methods &	2	
Results	Data Outputs	Statistical & Chemometric Outputs	• Basic statistical outputs (e.g., adj. p-values, standard deviations, test statistics) • Results of chemometric analyses (e.g., reported classifications/groupings of features or samples, observed trends in the data) • Visuals/plots (e.g., Venn diagrams, heatmaps, clustering dendrograms, volcano plots, network diagrams, PCA and loading plots) • New statistical metrics, algorithms, packages, and/or scripts	NA	
		Identification & Confidence Levels	• Reported identifications and associated confidence levels (e.g., levels described by Schymanski et al.) • Supporting data for annotation/identification (e.g., formula match scores, fine isotope pattern, retention time match, MS/MS match scores, source of MS/MS spectra) • For features with lower confidence IDs, (i.e., not standard-confirmed), proposed tentative structures and other annotated data • Semi-quantification or quantification data • Exported MS/MS spectra (e.g., as a library, database, or deposition into online repository)	3	
	QA/QC Metrics	Data Acquisition QA/QC	• Quality: Adherence to QA/QC protocols for sample preparation and data acquisition • Boundary: Description of the potential impacts of methods (sample prep, chromatographic, MS) on observable chemical space • Accuracy: Reported chromatographic and mass accuracy • Precision: Variability of observed retention time, precursor mass error, and abundance	1	
		Data Processing & Analysis QA/QC	• Quality: Outcomes of QC checks along the data processing & analysis workflow • Boundary: Impact of data processing & analysis method(s) on observed chemical space, observed limits of detection/ID • Accuracy: Performance measures (True Positive Rate, False Positive Rate, etc.) for known compounds or samples with known classification • Precision: Reproducibility/repeatability of performance measures for known compounds or samples with known classification; Calculations such as False Discovery Rate, F1 score, etc.	0	

Hyperlinked (HTML version) to supporting information

3-4 bullet point examples for each of the 13 sub-categories

Not exhaustive – intended to guide reviewers; relies on reviewer expertise/discretion.

Space for reviewer to explain assigned score

K. Peter, A. Phillips, et al. *submitted*

The Future of Exposure Monitoring

- High-quality targeted biomarker panels exist
  - Limitations on use based on study design
    - Spot sampling, cross-sectional design
  - Computational solutions to (partially) overcome limitations

- Massive data gaps given only targeted measurement panels
  - Most “known” chemicals have no monitoring data
  - Many previously “unknown” chemical stressors emerging

- Non-targeted methods emerging to fill gaps
  - Capture external and internal exposures/stressors
  - Issues with data quality and study reporting
    - Limits transparency & reproducibility
    - Data reuse not advised until quality and reporting issues resolved
      - Eventually will require semantic ontologies to enable comprehensive evaluations

BP4NTA Workgroup Leadership

Previous Co-Chairs:

Ben Place*
(NIST)

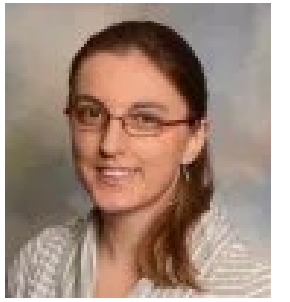
Elin Ulrich
(EPA)

SRT Committee Co-Chairs

Kathy Peter
(NIST)

Allison Phillips
(EPA)

Website Managers:



Sara Nason
(CAES)

Seth Newton
(EPA)

Current Co-Chairs:

Christine Fisher
(FDA)

Ruth Marfil-Vega
(Shimadzu)

Publications Committee Chair:

Natalia Soares Quinete
(FIU)

External Affiliations Committee Co-Chairs:

Tarun Anumol
(Agilent)

Elin Ulrich
(EPA)

*Current BP4NTA Documents Manager

EPA NTA Contributors

EPA ORD

Hussein Al-Ghoul*
Alex Chao
Jacqueline Bangma*
Matthew Boyce*
Kathie Dionisio
Louis Groff*
Jarod Grossman*
Chris Grulke
Kristin Isaacs
Sarah Laughlin*
Hannah Liberatore
Charles Lowe
Kamel Mansouri*
Aurelie Marcotte*
James McCord
Andrew McEachran*
Kelsey Miller
Jeff Minucci

EPA ORD (cont.)

Seth Newton
Grace Patlewicz
Allison Phillips
Katherine Phillips
Tom Purucker
Ann Richard
Randolph Singh*
Jon Sobus
Mark Strynar
Elin Ulrich
Ariel Wallace
John Wambaugh
Antony Williams

GDIT

Ilya Balabin
Tom Transue
Tommy Cathey

* = ORISE/ORAU

Questions?

sobus.jon@epa.gov

The views expressed in this presentation are those of the author and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.