Title: Personal Chemical Exposure informatics

Authors: Michael-Rock Goldsmith,* Christopher M. Grulke, Daniel T. Chang, Yue-Mei Tan, Raina D. Brooks,(2) Curtis C. Dary, Daniel Vallero

Affiliations: US-Environmental Protection Agency, Office of Research and Development, RTP NC 27711

1. US-Environmental Protection Agency, Office of Research and Development, RTP NC 27711

2. Student Services Contractor to the US-EPA, RTP NC 27711

***Corresponding author**: Michael-Rock Goldsmith
Mail Code: E205-01, 109 T.W. Alexander Drive, Research Triangle Park, NC 27711
Phone: 919-541-0497
Email: goldsmith.rocky@epa.gov

***Disclaimer:** The United States Environmental Protection Agency has provided administrative review and has approved for publication. The views expressed in this chapter are those of the authors and do not necessarily reflect the views or policies of the United States Environmental Protection Agency.*

Keywords (4-5): chemical exposure, personal exposure, nearfield exposure, consumer product ingredients,

@epa_gov's Michael Rock Goldsmith: "If dose makes the poison, and exposure makes the dose, then exposure also makes the poison" #GlobalChem 10:22 p.m. Tue, Mar 6

Chemical Exposure science is the study of human contact with chemicals (from manufacturing facilities, everyday products, waste) occurring in their environments and advances knowledge of the mechanisms and dynamics of events that cause or prevent adverse health outcomes. (adapted from United States National Research Council (2012). Exposure Science in the 21st Century: A Vision and A Strategy, p. 19.) **Personal chemical exposure** (PCE) is a subtopic within exposure science that focuses on the individual (see Figure 1. A). PCE primarily look at the following questions: "Who you are?" (the receptor), "What you are doing?" (activity), and "What do you interact with or use in what you are doing" (products, articles, feeding, exercising, hygiene, etc…). Mathematically (deterministically or probabilistically, or both) estimating a *receptor's* (target species such as a human) chemical exposure is also achievable by knowing "where" you are (location), and "when" you are there (time), as it contextually limits the choices of "what you are doing" into a tractable probability. Chemical exposure, in general, is a function not only of the chemical, but of the individual receptor's biology, environment and life-style factors that make up an individuals "personal experience" (See Figure 1. B). By combining individual habits with chemical residue or composition knowledge a full picture of that individual's personal chemical exposure can be obtained.

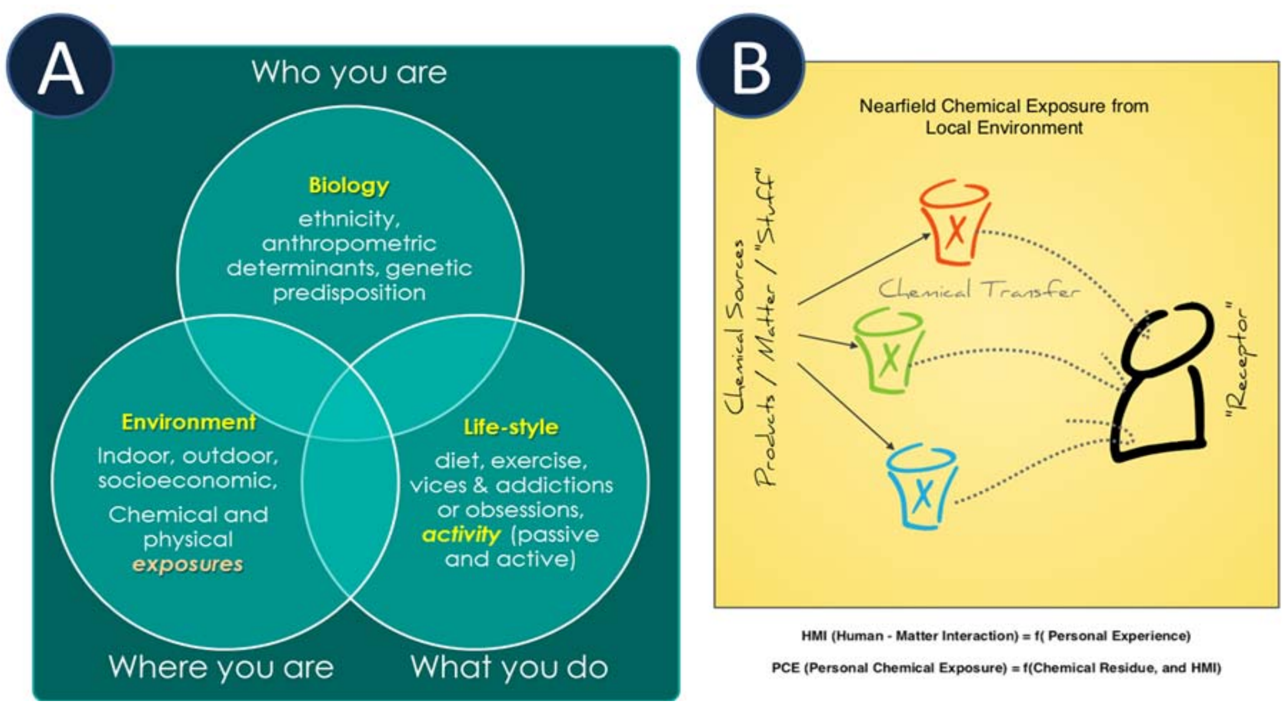

Figure 1: A) *The three information domains that directly affect personal chemical exposure related to "Who? What? Where?" of an individual* ***(B)*** *Nearfield chemical exposure from one's local environment results as chemical transfer between various sources of chemicals (red, green and blue cups of idealized chemicals X), and is a function of how receptors (humans) interact with products, materials, or articles (Matter) and the chemical residue constituents of those products.*

When coupled with chemical-specific hazard characterized by known dose-response relationships (e.g., toxicological characterization, Adverse Outcome Pathways (AOPs), exposure thresholds), **personal chemical exposure** holds the potential to estimate personalized (or Individual) risk by bringing exposure into the "risk mix". A visual overview of the published literature related to personal chemical exposure reveals information about chemical exposure sources, exposure routes and scientific methods and data that characterize personal exposure more generally is shown in Figure 2.

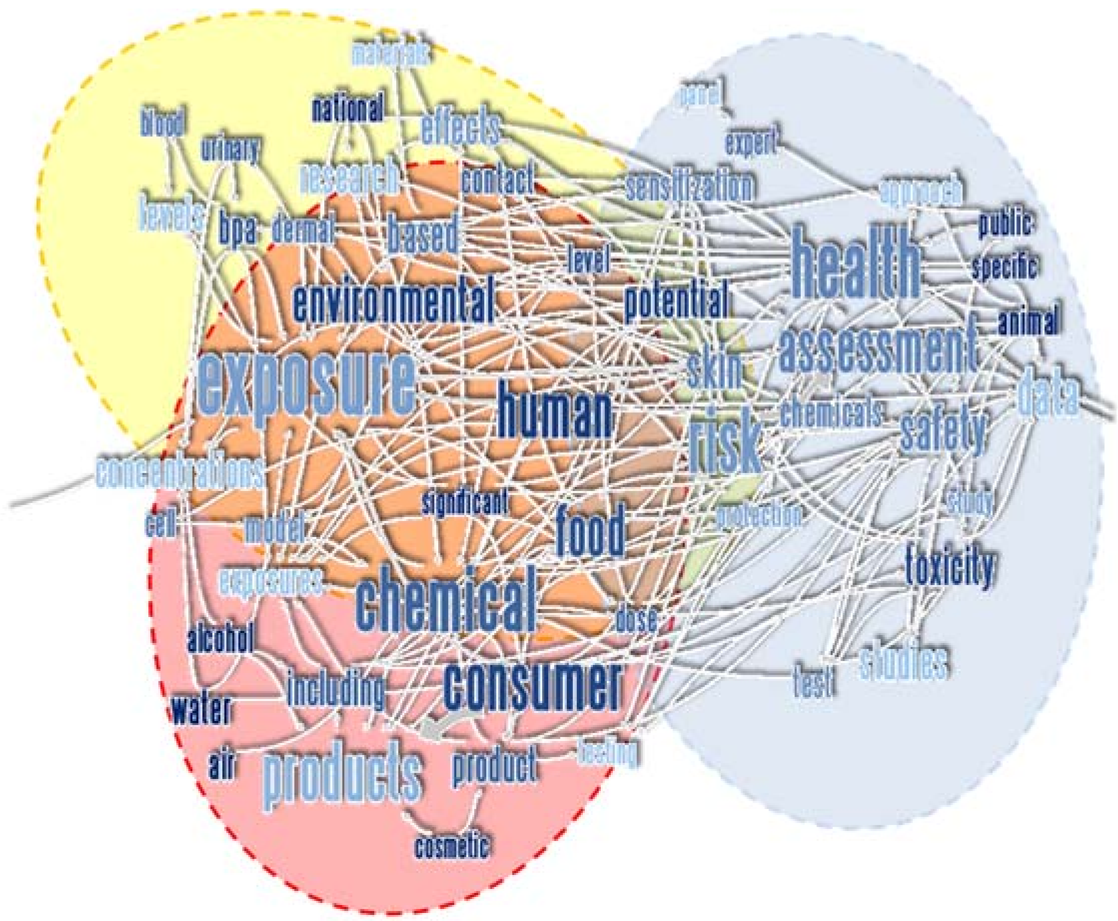

Figure 2: A text-mining visualization of "Personal Chemical Exposure" according to the literature: Pubmed search for: ("everyday exposure" OR "consumer exposure" OR (("consumer product" OR "consumer products") AND exposure)) AND chemical) with advanced filters: "Humans" provides 224 research articles (05/23/13). The resulting text abstracts were imported into Many-eyes.com and visualized using a "Phrasenet" analysis. Key themes are circled in colored ovals related to (a: Red)) sources of everyday exposures (b: Yellow) exposure routes for these everyday exposures and (c: Blue) key areas of research involved in safety assessment. The shortened URL for the underlying data is http://goo.gl/qSjWe .

Personal Chemical Exposure Informatics (PCEI) is the sum or combination of all informatics resources required to make personal chemical exposure extrapolations (via modeling and simulation approaches) at the crossroads of a biological receptor with the chemical constituents of its interactome or exposome across space and time. Moreover, in its more generalized analog, personal informatics systems, we know that "people want to get information about themselves to reflect on, and that systems that support this activity need to be effective and simple to use." The nature of the data PCEI tools need to handle is diverse and ranges from chemical, biological, social, and geographical data types, including but not limited to:

- Air quality (indoor/outdoor),
- pollution (for instance listed on TSRI toxic substances release inventory),
- location (GPS data) ,
- time activity journals (i.e. American Time Use Survey / Comprehensive Human Activity DB),
- product use in those activities (Exposure factors such as the Exposure Factors Handbook)
- material composition (articles and consumables, i.e. National Library of Medicine's Household Product Database),
- water quality (for instance the ENvironmental Working Group's National Drinking Water Database : http://www.ewg.org/tap-water/)
- anthropometric determinants
    - species, life stage, gender, ethnicity or strain, underlying physiological/genomic data
    - dietary, metabolic output / physical activity …
    - chemical and species-specific ADME (pharmacokinetics or Absorption, Distribution, Metabolism and Elimination)

In addition, PCEI needs to allow collection and correlation of proxies that allow one to deduce or infer these variables. The complexity of PCEI needs an information model that contains all the required data sources and proxies necessary to model personal chemical exposure. This information model and required data-streams to extrapolate personal chemical exposure (PCE) are shown in **Figure 3**. Specifically, the aim of PCEI is to better understand the exposure sources present in micro-environments; some of these efforts have been explored through a proof-of-concept study for the first step of Systems Reality Modeling – Chemical Inventory (http://www.systemsreality.org). Despite its preeminence in the source-exposure dose-effect continuum, source (chemical source) information is generally not well integrated in exposure and risk assessment. In exposure studies, source information is often collected in an ad hoc fashion either by the investigators attempting to inventory products in homes or offices; or by the study participants attempting to interpret chemical information from consumer product labels. An approach that will significantly increase the collection efficiency, amount, and quality of source data one can collect in a short period of time is outlined in **Figure 3-5**. This approach holds the potential to reduce the uncertainty and errors in real time data acquisition and interpretation.

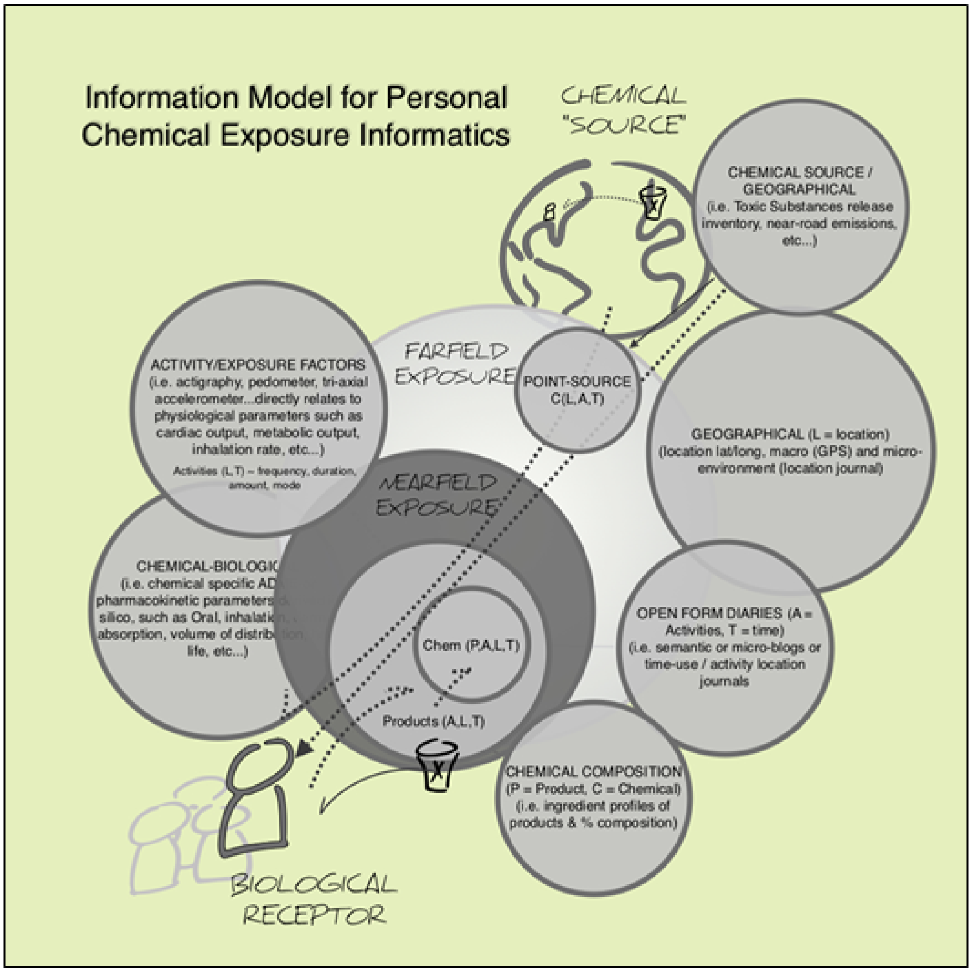

Figure 3: *Information model for Personal Chemical Exposure Informatics. The information model contains all requisite information for calculating chemical transfer from a chemical source to a biological receptor (an individual). Legend: C = Chemical, P = Product, A = Activity, T = time, L = Location,. Depending on the individual biological receptor, and the activities and locations of those individuals they will have a unique set of activities they perform, with a corresponding set of products with known chemical formulations and specific chemical exposures. The exposure profile of an individual is unique, and at the aggregate level there may be some community/population based similarities, or even occupational related chemical exposures that can be deduced using such an approach.*

Integrating Personal time-activity use into PCEI: From information model to workflow

While it is vital for exposure modelling to gather current time-location activity profiles of the U.S. population to better understand relevant routes of exposure, these profiles are difficult to obtain. The state-of-the-science in gathering these data is time-consuming, expensive and subject to a number of limitations. It places a great burden on participants to provide data, leading to noncompliance with study protocols, biases in activity reporting, or even participant dropout. Thus, to increase data collection efficiency and simultaneously reduce participants' burden of filling out conventional time-activity diaries, an innovative approach that monitors individuals anonymously in real-time without solicitation is much needed.

While seemingly far-fetched, a growing subset of the population actually volunteers their time-activity information to the public via social media (e.g., "tweets", "status updates", "check-ins") making the development of a tool for converting such free form texts into a structured set of "knowledge" ideal. (Although the sole reliance of information from social media may lead to limited scope and possible bias.) Thus, developing an "activity tracking app" for smartphones to reduce the burden on participants in traditional observational studies is vital for PCEI. Such apps would have a great potential to collect real-time activity data with high spatial and temporal resolution when installed on a smartphone with other "monitoring apps" such as GPS for tracking location, accelerometer for tracking movement, calorie-tracker for recording food consumption, and time-logger for recording time-activity information. While social media and other available data streams (e.g., geospatial location) can potentially provide a wealth of untapped information to mine human activity data, the challenge of separating the noise from relevant data will also need to be dealt with. However, as a minimum, developing preliminary "tweet/micro-blog" codification algorithm which can be further optimized by including knowledge of location/activity co-occurrence is also needed for PCEI. Such "passive interrogation" methods show a great deal of promise for increasing the quality and quantity of data to inform exposure, epidemiology, and behavioral sciences with minimum participant burden.
With an eye towards mobile computing as the desired computational environment to deploy a Personal Chemical Exposure Informatics we have designed a workflow that captures the interactions of an individual with their local environment (**Figure 4 From A to B**) and allows an individual the ability to explore their personal chemical exposures.

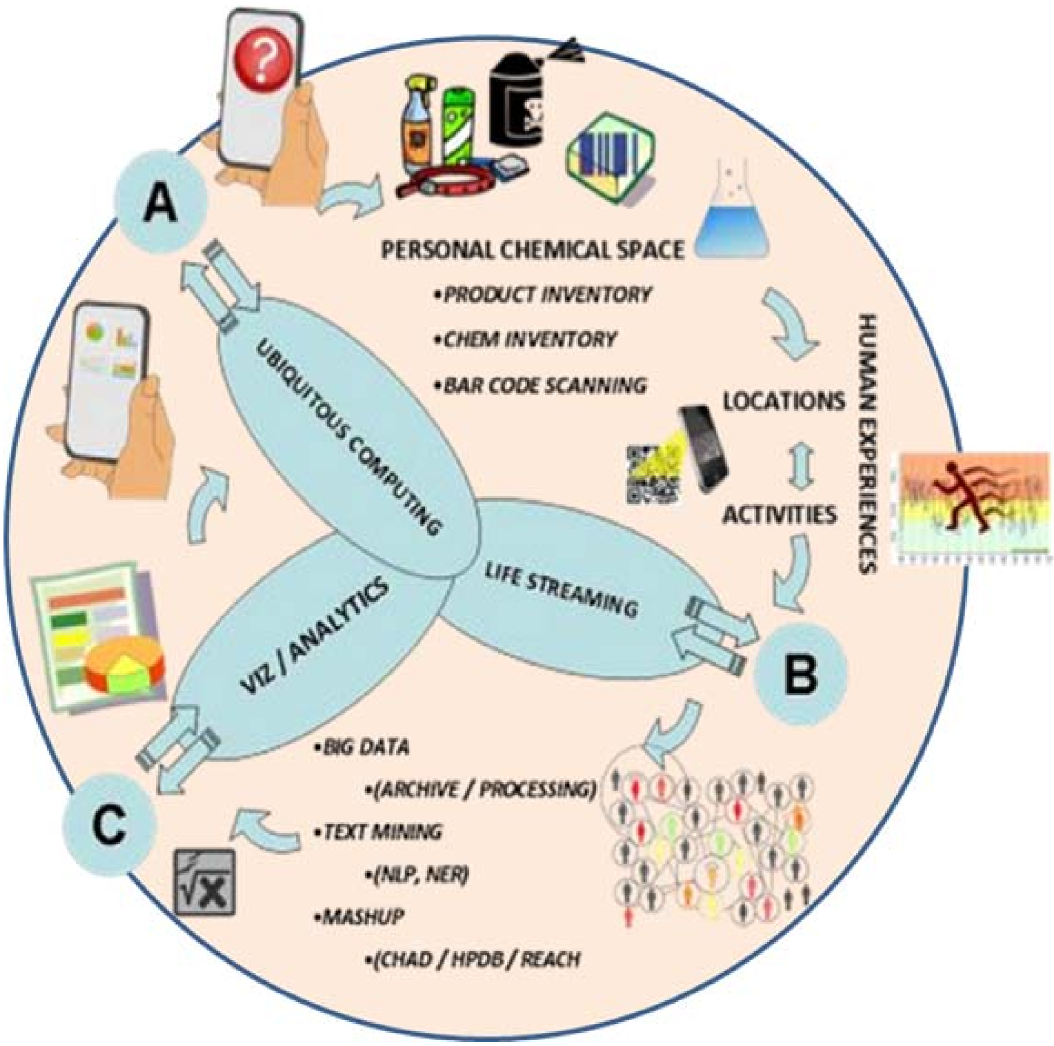

Figure 4: *To inform human matter interaction and personal chemical exposure we propose the use of a PCEI workflow (a) Ubiquitous computing (B) Life-streaming and appropriate data-mining and (c) Vizual analytics and modeling / mashups to make use of big data and inform the individual or personal chemical exposures from everyday household products. NLP = Natural Language Processing, NER = Named Entity Recognition, CHAD is the Comprehensive Human Activity Database (http://www.epa.gov/heasd/chad.html) , HPDB is the Household Product Database and REACH is the European CHemical REgulation program "Registration, Evaluation, Authorisation and Restriction of Chemical substances" .*

Integrating the disparate "information pieces" into a PCEI product for smartphone Apps

Ideally, integrating the disparate data streams (Figure 3 geographical chemical release information, geographical position data, exposure factors, activity logs, chemical composition and uptake/intake) into a mashup which uses and combines this data will allow for real-time exposure assessment based on a particular real time scenario. Within the a local environment, personal chemical space – bar coded products associated with an inventory which links concentrations and usage – could be linked with location and activity information to determine exposure based on contact with the chemical substances within the products. The exposure assessment for particular products could then be informed by mined text mapped to archived "Big Data" containing activities, frequencies and durations mapped to typical product usage information from CHAD, the HPDB, SPIN and REACH. A computational evaluation of exposure will be determined for each given product. Finallythese analytics could be relayed back to the user for each product scanned in their personal space. This combination of visualization and aggregation would generate a user-friendly presentation of the exposure assessment thereby increasing the functionality of the data streams for the user; A user could then have an accurate representation of the summative chemical exposure potential of their local environment.

Using the information models, PCEI workflow, and user interface (UI) one can conceivably develop a research tool (either smartphone app or desktop / web application) that allows for simulation of exposures within local environments. The user can specify day to day activities in a series of scenarios and obtain a virtual exposure assessment drawing on information from the mashed-up data streams. These simulations will allow for the exploration of potential exposures on virtual subjects in a myriad of microenvironments without the limitations of obtaining new location, activity and chemical specific data.

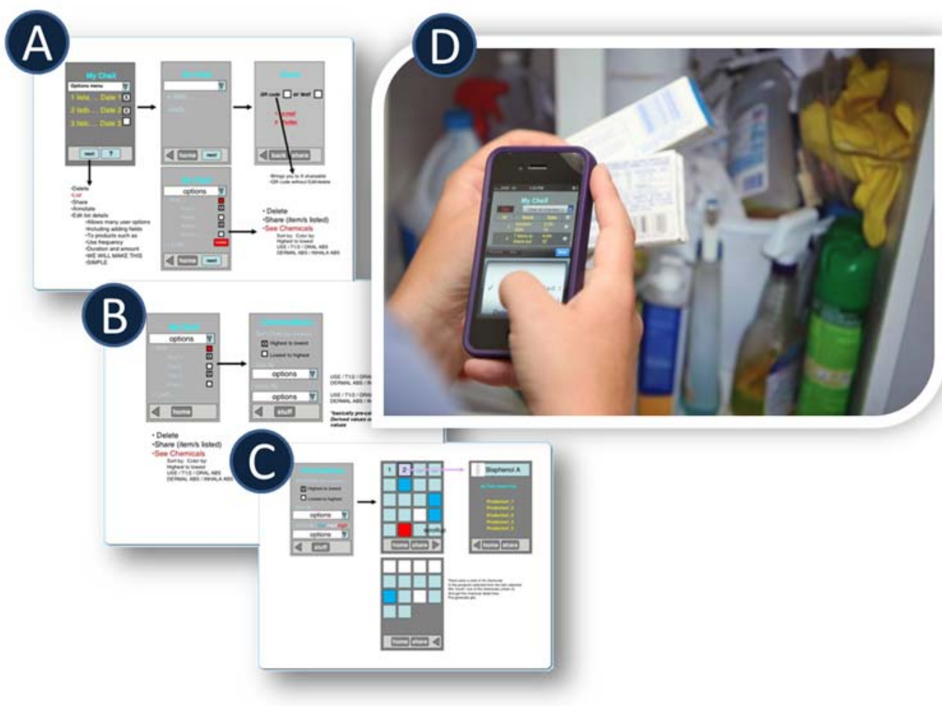

Figure 5: The designed mock-up of a smart-phone User Interface (UI) for (A) entering a chemical inventory, with options to select and share subset "product" lists (b) create sub-lists and identify how to visualize the chemical exposure by a variety of sorting and coloring schemes such as half-life or absorption and (C) a sharable tactile link table with colors according to desired scheme that links each chemical to toxicity resources such as ACToR (http://actor.epa.gov) (D) a real "prototype" in the proper contextual setting....at home.

Modeling and simulation using data derived from PCEI increases the predictive capacity of individuals, researchers, industries and regulators because it can be used to extrapolate exposure information to new chemicals. The simulation models can include visualization underlined with mathematical models describing the exposure processes from source-to-receptor to calculate the impact of product usage on chemical exposure which can be linked to health effects and subsequently inform decisions to mitigate potentially harmful exposures or to promote more sustainable health decisions.

References (3-5)

1. Summary Report for Personal Chemical Exposure Informatics: Visualization and Exploratory Research in Simulations and Systems (PerCEIVERS) June 26,27- 2012, **EPA/600/R13/041**

2. Systems Reality Modeling: an EPA Pathfinder Innovation Project. http://www.epa.gov/heasd/research/srm.html

3. Aggregated Computational Toxicology Resource ACToR. http://actor.epa.gov

4. Li, I., Dey, A., et al. A stage-based model of personal informatics systems, In Proc. CHI 2002, ACM Press (2010).

Additional off-line reading (4-5):

1. Reeves, D. R., and Guinn, C. I. Improving Upon Semantic Classification of Spoken Diary Entries Using Pragmatic Context Information. In Proc. IC-AI 2008, CSREA Press (2008).

2. Yau, N. and Schneider, J. Self Surveillance. Bulletin of the American Society for Information Science and Technology 35, 5 (2009), 24-30.

On-line resources

1. The Systems Reality Modeling landing page - http://www.epa.gov/heasd/research/srm.html

2. PerCEIVERS workshop - Personal Chemical Exposure Informatics - Visualization user experience, systems modeling and simulation research - http://www.epa.gov/ncct/expocast/percei.html

3. http://www.good-guide.com

4. The Aggregated Computational Toxicology Resource - http://actor.epa.gov

5. National Library of Medicine's Household Product Database (HPDB) http://householdproducts.nlm.nih.gov/

6. Consumer Product Information Database. http://whatsinproducts.com/