Title: PAVA: Physiological and Anatomical Visual Analytics for mapping of tissue-specific concentration and time-course data

Authors: Michael-Rock Goldsmith, Thomas R. Transue, Daniel T. Chang, Rogelio-Tornero Velez, Michael S. Breen, Curtis C. Dary

Affiliations:

1. National Exposure Research Laboratory, Human Exposure and Atmospheric Science Division, U.S. Environmental Protection Agency, 109 T.W. Alexander Drive, Research Triangle Park NC 27711
2. Lockheed-Martin Information Technology, A Contractor to the U.S. Environmental Protection Agency, 109 T.W. Alexander Drive, Research Triangle Park NC 27711

Abstract: We describe the development and implementation of a Physiological and Anatomical Visual Analytics tool (PAVA), a web browser-based application, used to visualize experimental/simulated chemical time-course data (dosimetry), epidemiological data and Physiologically-Annotated Data (PAD). Using continuous color mapping scheme both spatial (organ shape and location) and temporal (time-course / kinetics) data was cast onto an abstract, layered, 2D visual representation of the human anatomy and physiology. This approach is aligned with the compartment-level of detail afforded by Physiologically-Based Pharmacokinetic (PBPK) modeling of chemical disposition. In this tutorial we provide several illustrative examples of how PAVA may be applied: (1) visualization of multiple organ/tissue simulated dosimetry of a previously published oral exposure route ethanol PBPK model, (2) visualization of PAD such as organ-specific disease time-lines or (3) tissue-specific mRNA expression-level profiles (e.g. phase I / II metabolic enzymes and nuclear receptors) to draw much needed molecular biological conclusions at organ-level resolution conducive to model development.. Furthermore, discussion is raised on how graphical representations of PBPK models, and the use of PAVA more generally to visualize PAD, can be of benefit. We believe this novel platform-independent tool for visualizing PAD on physiologically-relevant representations of human anatomy will become a valuable visual analytic addition to the tool-kits of modern exposure scientists, computational biologists, toxicologists, biochemists, molecular biologists, epidemiologists and pathologists alike in visually translating, representing and mining complex PAD relationships required to understand systems biology or manage chemical risk.

Keywords: physiologically-annotated data; dosimetry; visualization; visual analytics; anatomical; physiological; server-side application; model animation; concentration time-course; disease progression timelines; model rendering; PBPK; PBTK

Introduction

"A picture is worth a thousand charts" (1)

The need to both more efficiently analyze animal and human chemical disposition data derived from *in vivo* animal dosimetry studies or *in silico* physiologically-based pharmacokinetic (PBPK) modeling efforts and the ability to see complex disposition data in a rational and clearly engaging manner has created a unique opportunity for visual analytics applied to physiologically-annotated data (PAD). To date, efforts to anatomically and physiologically map tissue dosimetry (chemical time-course) data have been limited to fish and rat models. (2,3) Although such models could be modernized in terms of required data-feeds and software packages (and in some instances antiquated architecture) none currently support an open-source third-party requirement or an open-access platform. Similarly, despite the existence of several open-access social visual analytics initiatives available on the internet for visualizing data, none have dealt with rendering pharmaco / toxico-informatics in the context of human anatomy.(4)

Both the abundance of chemical disposition data in the literature, coupled to the availability of open-access PBPK modeling packages (see Table 1) has engaged a much larger scientific community, making the task of developing and/or applying PBPK models in the context of chemical risk management more globally accessible for both therapeutic and environmental chemicals in regulatory and industrial settings alike. A major motivation for modeling efforts that necessitates the translation of exposure to internalized dose in conjunction with mechanistic toxicity data, is in support of quantitatively informing the dose-response relationship, or more formally the toxicological paradigm that "the dose makes the poison".

Table 1: Open-Access Physiologically-Based Pharmacokinetic modeling software packages and resources available on the world-wide-web.

Open-Access PBPK Software	Uniform Resource Locator	Citation
ERDEM	http://www.epa.gov/heasd/products/erdem/erdem.htm	**(5)**
PKQuest	http://www.pkquest.com	**(6,7)**
BOOMER	http://www.boomer.org	**(8)**
PBPK 1.0 for Excel	http://www.trentu.ca/academic/aminss/envmodel/models/PBPK.html	**(9)**

Similarly, other organ/tissue-specific data such as tissue-specific pathology (e.g. organ-specific cancer timelines), protein expression-level data and organ/tissue chemical composition data could provide additional streams of data for which anatomically mapped visualization are amenable and would complement both modern methods of managing chemical risk and fundamental chemical/biological research alike. We believe that the very diversity and complexity of PAD (e.g. *in vitro*, *in vivo*, *in silico*, histopathological, and epidemiological) married to modern hypothesis-driven scientific inquiry requires the consideration of novel ways to visualize such data. These novel visual analytics methods cast on anatomically and physiologically relevant representations of the "human system" should prove to be fundamentally informative and useful.

In the process of developing a method for physiological and anatomical visual analytics (PAVA) we have outlined critical rationale for why one would want to map dose visually in an image/movie/animation format, with several motivating factors including, but not limited to:

1. Potentially easier target ***identifiability*** with relation to actual organ/tissue constructs and less abstract than multi-series time-course plots.

2. More ***compact*** and portable format to visualize, since anatomically / physiologically mapped animations occupy a single figure instead of multiple plots. Justifiably, the more compact the visualization is, the "greener" it is as well.
3. Easier to visualize ***comparative*** pharmacokinetics on a common physiology: subtracting time course data between (a) different chemicals (b) parent/metabolite (c) life-stage differences (d) physiological variability (e.g.. BMI, body weight) (e) gender differences
4. Arguably easier to ***relate***, clinically, to site of relevant doses by providing a visual map of anatomical and physiological acquisition of dose over-time: Clear localization of critical organs to identify for clinical pharmacokinetics, such as those found in radio labeling studies.
5. No ***linguistic barriers*** for interpretation of data (as long as the individual knows anatomy and physiology or how to identify some basic organs, the information is conveniently capitulated graphically.
6. ***Pedagogically*** tractable tool for visual learning that provides engaging clarity to both scientific investigators for interpretation of large complex datasets, or for translating complex physiological relationships to non-scientists.

In this spirit we have developed a clear yet comprehensive abstract layered road map of the human anatomy, simplified for representing physiologically-annotated data (PAD) (e.g. tissue dosimetry, or even tissue-specific disease metrics) that can be used in conjunction with any of the aforementioned open-access packages or proprietary packages alike.

Method

Developing an abstract, layered, 2D image-map visual representation of anatomy and physiology

Although there has been headway in the fields of radiology for visualizing dose on computed tomography 3D representations of human anatomy such as the Medical Internal Radiation Dose (MIRD) Committee's volumetric pixel (VOXEL) methods, the level of detail found in most PBPK models are considerably more coarse-grained than the sub-organ/tissue level of resolution**.**(10) For this reason a readily generated and rarefied

graphical representation of the human anatomy and physiology was used to develop PAVA. Additionally, PAVA design was focused on platform and software independence with accessibility for soft computing and data visualizing by internal scientists. For this reason we had chosen to develop a web-based application (CGI/perl Script in HTML) server-side application that would make use of a layered image for each independent organ and would map dose through color onto said tissue/organ.

Our model was produced by collecting a series of image layers, each representing one organ or tissue and aligned with a background layer representing the human figure from thigh up. For the HTML mock-up these are loaded as separate images and their "visibility" style (cascading style sheet, CSS) attribute switched between "visible" and "hidden" to toggle them on or off using a client side JavaScript. For static or animated graphics of the body plus organs/tissues, a CGI-based PERL script calls *ImageMagick* software to first recolor each organ/tissue, then assemble relevant images to form a figure. (11) Finally, a time point label is added and multiple figures are assembled for animationsFor the current version, the selection of organs to consider was based primarily on the observation of their occurrence as model tissue compartments found in current literature PBPK models and does not imply that more subtle delineation of individual physiological constructs should not be represented in future versions. An overview of the anatomy used in PAVA is shown in Figure 1. Future versions may include more organs/tissues as well as more clearly depicted or better-separated graphics. For clarity, only a subset of some tissues or organs is depicted. For instance, the representation of gender-specific reproductive organs was combined (i.e. uterus/ovaries)

and prostate/testes, though future versions of PAVA may include scalable vector graphics (svg) representations of substantially more organs.

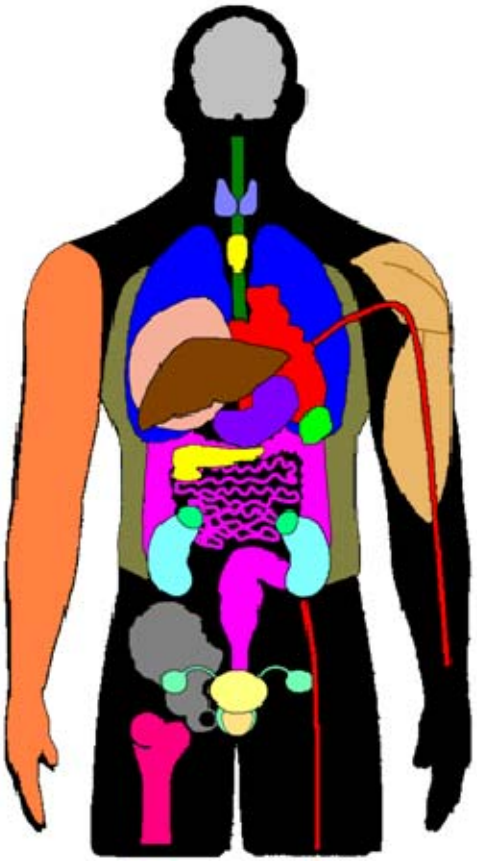

Figure 1. The organs we have rendered in PAVA are adipose, bladder, bone, brain, gonads, esophagus, bone marrow, blood, adrenal, heart, kidney, intestines, lung, liver, muscle, pancreas, prostate, mammary, spleen, thymus, thyroid, stomach, skin and "rest-of-body"/carcass. The color scheme here is merely to delineate differences in organs.

Visual mapping and representation of Dose: $Dose_{max}$, $Dose_{mid}$ and $Dose_{min}$

Individual dose representation of each organ is achieved by normalizing the tissue-specific concentration data to the maximum concentration across all compartments. Relative dose values are represented according to the following color scale - white for zero dose ($Dose_{mid}$), red for maximum dose ($Dose_{max}$) and an increment between zero

dose to max dose as a continuous gradient between white and red. Similarly, a "negative" dose ($Dose_{min}$, achieved when comparating two pharmacokinetic profiles or PAD models resulting in negative tissue specific values compared to a reference profile) can be mapped in blue for comparative purposes.

PAVA script development and interface.

Although the interface for PAVA is standard HTML, the back-end is comprised of both CGI and Perl scripts that handle several tasks related to (1) data submission and transformation (2) tabular (matrix) dose rendering and (3) animation (animated-GIF) generator. The PAVA interface has the following user-specific features (See Figure 2):

- Gender selection menu (drop-down list)
- Time-step and start-time entry
- Organ-specific disposition table from output, text box entry
- Organ selection menu with order sorting
- Real-time updated 2D anatomical representation based on organ selection
- GO! Button (submission of data to the CGI script)

After setup and submission of data to the CGI script data is rendered in a table format with an anatomical representation of the defined organs, in addition to a physiological representation (individual organs near each row of time-course data). Finally, the option to render an animated GIF by selecting the length of the animation (seconds) and number of frames-per-second that interpolates dose linearly between time-steps is used.

Illustration of Anatomical Tissue Dosimetry Rendering using PAVA: Model output of oral ethanol fat diet regimen

For the PAVA output in this paper, the PKQuest java applet PBPK modeling package was used with the pre-existing fatty-diet oral ethanol ingestion models.(7) The default output of PKQuest was an excel-formatted output file containing tissue specific time-course data (column: organ, row: time-course) that we had selected.

The model output data was pasted into PAVA (see Figure 2), organs were selected in order of occurrence in the table using the graphical organ selection tool, time step was selected in appropriate units, and the resulting graphical anatomical version of organs were updated in real-time by our anatomical layout tool. Next, we submitted (select Go! Button) the input table and obtained a table representation of our model with individual cells colored according to a gradient system scheme (white – 0 dose, red = maximum organ dose, all other concentrations normalized and represented on the white-red gradient). Finally, by selecting the length of the subsequent animation (in seconds) and number of frames-per-second to linearly interpolate for smooth-graphics rendering, multiple frames were generated and concatenated into a single animated-gif file rendered in *ImageMagick* running in the background. The resulting PAVA animated graphic updated in a new web-browser.

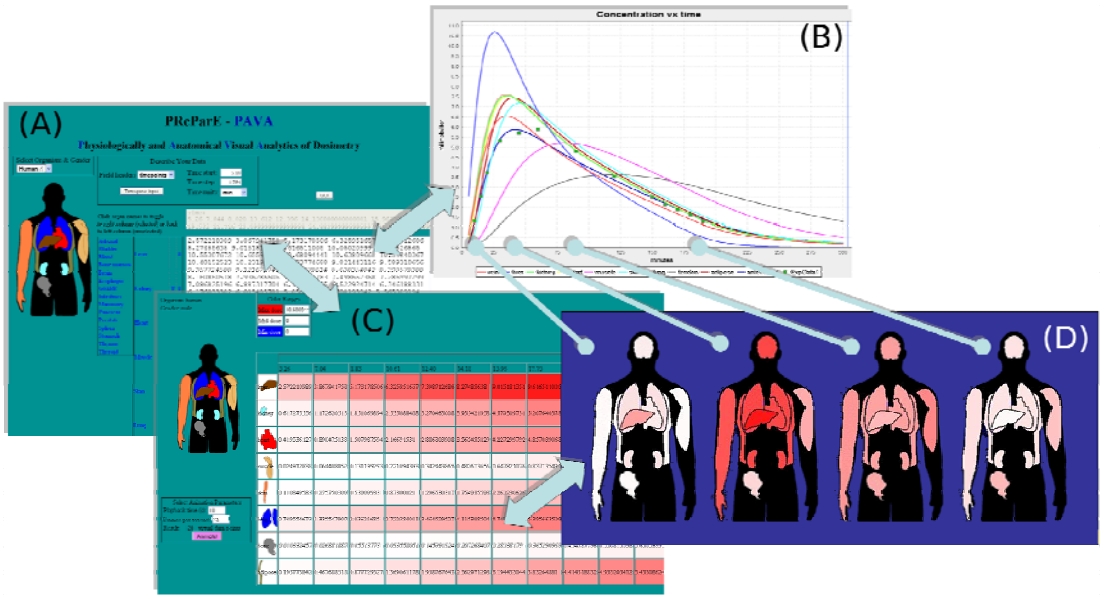

Figure 2: PAVA-rendered PBPK model simulation output for ethanol from PKQuest.(7) (A) raw data and organ selection as shown in the PAVA input screen and (B) output data as rendered in plot fashion from the model simulation environment (PKQuest) versus (C) matrix representation with individual physiological components and (D) full time-course animation. The animated GIF in (D) reveals a clear cut lag rapid distribution to all tissues except for muscle (a little slower) however clearance in other tissues appears to occur faster than muscle. Visualizing this data on the human form can sometimes bring attention to details not typically seen in standard graph layout. For instance, one may speculate from visualization of this ethanol PBPK model that as a result of muscle tissue intoxication (i.e. dose to muscle) may give rise to impaired motor-skills even after reduced brain or blood levels.

PAVA rendering of Tissue-Specific mRNA Expression of oxidoreductases versus carboxylesterases versus all other Phase II enzymes across eight (8) human tissues

Tissue-specific mRNA expression data for nuclear receptors (NR) and cytochrome P450 (CYP) and phase II metabolism enzymes have been taken from the literature and summed over all biomolecular classes across all sub-types.(12, 13, 14) Furthermore, carboxylesterases (hCE1, hCE2 and hCLE) were separated, also making a critical component of phase I metabolism. Finally the individual metabolism classes (NRs, CYP450, EST, PHASE2) were normalized against each of their maximum values across all tissue space. We have chosen to show only those tissues in which all classes have been measured.

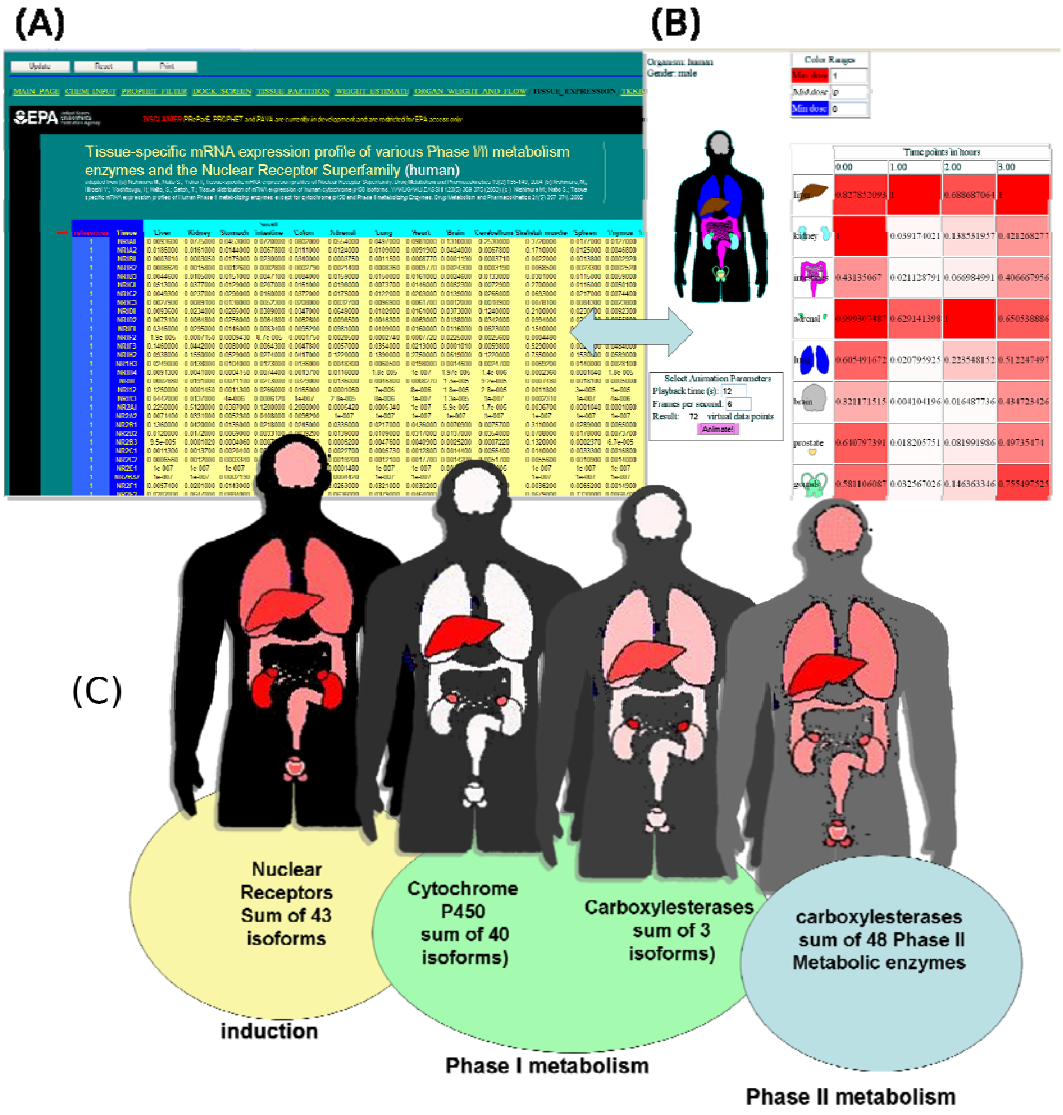

Figure 3: (A): Raw data tissue-specific mRNA expression profile (12, 13, 14) rendered anatomically with PAVA across (I) Nuclear Receptor (NR), (II) Cytochrome P450 (CYP450) (III) esterase (CES1, CES2, CEL) and (IV) 97 additional Phase 2 enzymes (including but not limited to STAs, UGTs, TAs, etc…). (B) physiological and matrix view of tissue-specific expression profiles and (C) anatomical representation. What can be observed is that although nuclear receptors (yellow circle) and phase II metabolism enzymes (blue circle) are expressed in all tissues rendered, both the oxidoreductases (Cytochrome P450 sum of 40 isoforms) appear to be centralized in liver and adrenals whereas esterase activity appears to be centralized in liver, intestines, and to a lesser extent, lungs (Cytochrome P450 and Carboxylesterases in green circle).

PAVA rendering of Tissue-Specific Cancer Mortality Data in multiple human tissues between Men and Women over the period of 1930-2000

Tissue-specific cancer death rates among men and women in the USA over the were used in PAVA to visualize historical organ-specific disease-state progression over a 70 year period (1930-2000), which were age-adjusted to the 2000 US standard population.(15) The PAVA animations shown in Figure 4 for (a) men (b) women and (c) man – woman difference timeline are shown to illustrate the analysis possibilities in PAVA in visualizing disease mortality or incidence data

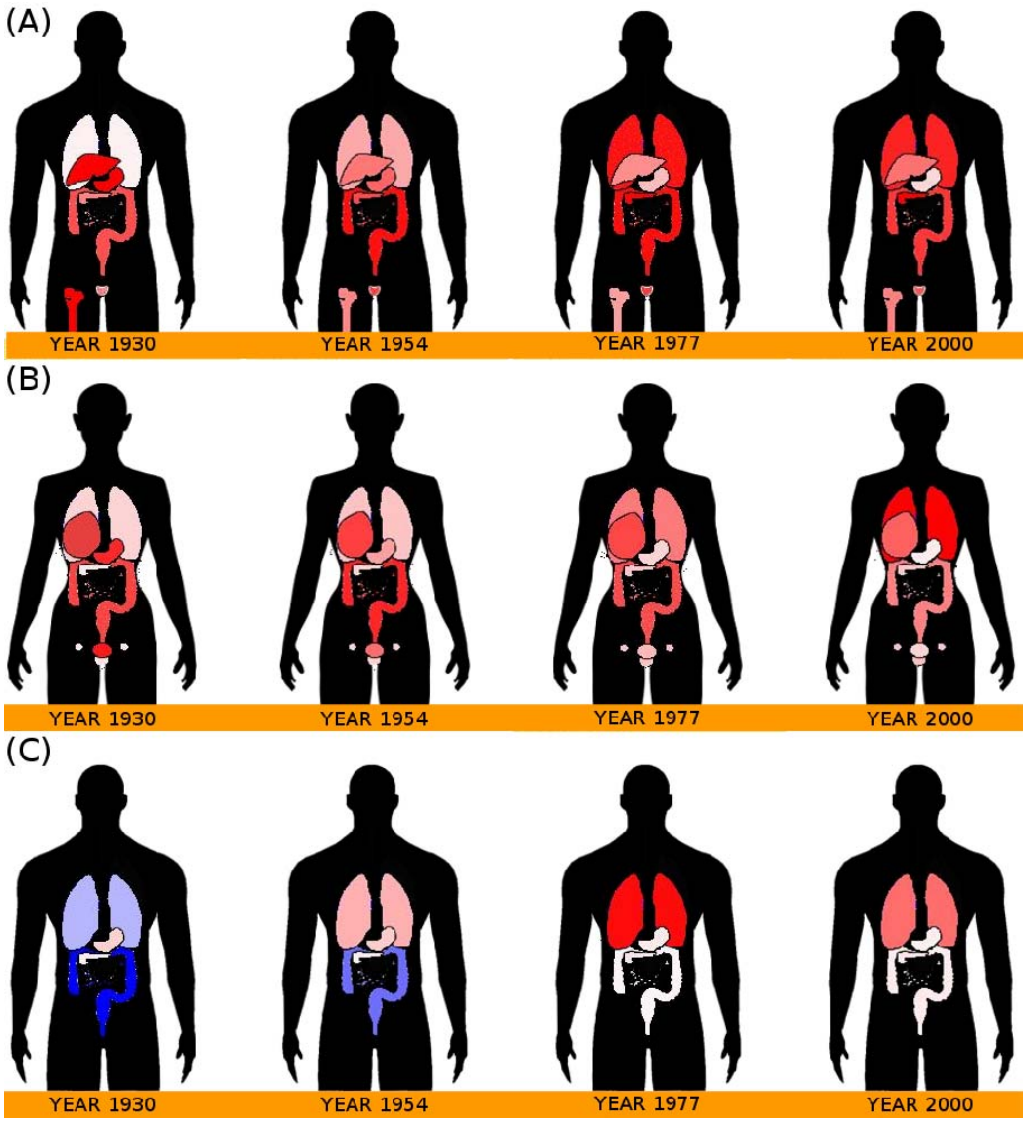

Figure 4 Tissue-specific cancer mortality rates among American **(A)** men **(B)** women or **(C)** Difference prevalence maps between men and women (MEN - WOMEN) for select organ-annotated cancer types over the period of 1930-2000. For **(A,B)** intensity of red is normalized to highest tissue-specific cancer rate over entire 70 years whereas for **(C)** Red indicates positive in males, blue indicates positive in females, white indicates similarity and intensity indicates magnitude of difference.

Discussion & Conclusion

We have demonstrated that visual representation of data from individual tissues can provide an interesting, compact, and informative way of representing simulation and

pharmaco / toxicokinetic data. Three specific applications were highlighted in the current development of PAVA – I. visual mapping of a PBPK output with tissue dosimetry and II. visual representation of tissue-specific expression-level data as they pertain to chemical interactions (e.g., nuclear receptors, phase I and II metabolic enzymes) and III. reporting tissue-specific disease mortality (epidemiological) timelines.

The animated GIF of ethanol consumption (Figure 2) revealing the residual lag-time in muscle was clearly observed and easy to ascribe dose to tissue. From a clinical standpoint it would provide a visual map for tissue assay selection at a given time with no linguistic barrier; a visual representation that clearly any clinician could comprehend.

Similarly, our example of displaying tissue-specific expression-level data (Figure 3) provides information as to how rarefied one may or may not want to develop a PBPK model. For instance, if a chemical agent was known to interact with any one, or several, of each of the metabolic enzymes and/or nuclear receptor subclasses, it would be useful to consider developing models that included as many tissues as possible where expression of putative biomolecular interactions may occur. As an example, if one desired to track the fate of a chemical which, *in vitro,* underwent predominantly enzyme-mediated ester hydrolysis, an intestinal compartment as well as a liver compartment and kinetics for each of those compartments should be explicitly considered, as might have been inferred from relevant data aggregation and visualization efforts. Similarly, for subsequent phase II metabolism of such a chemical one would select or potentially rarefy the model to include more tissues since phase II enzymes appear to be expressed more uniformly in multiple tissues.

Finally using tissue-specific mortality PAD (Figure 4) provides an interesting visual time-line that clearly shows progression of cancer-mortality types with an overlay of at least oral and respiratory exposure route related tissues (lung & bronchus, stomach, colon/rectum) that may or may not assist in historically identifying trends in social chemical exposures (environmental, therapeutic, and life-style) that may give rise to such disease incidence. Interestingly, looking at differences between man and woman for four common tissues may lead an investigator to hypothesis generation and speculation required for rational scientific progress. For instance one might speculate:

- Is the higher level of female colon cancer from 1930 – 1950 (blue) a result of different dietary, life-style, chemical occupational exposure from working in factories?
- Is the apparent loss in female GI related cancer a result of the onset of a widely adapted life-style chemicals, environmental exposures or dietary changes?
- Has the prevalence of male lung cancer been ascribed to a male-specific or dominant activity?(See Figure 4A and 4C) from 1954 onwards. Is the lag seen in females (Figure 4B) and difference between men and women (Figure 4C) decreasing in more recent years as gender bias alleviates and life-style choices of women become closer to men (2000 difference map is lighter pink than 1977)?
- Is the apparent similarity between pancreatic, stomach and colon cancer possibly not related to a gender-biased activity?
- Based on the male timeline: the reduction in stomach, colon and blood related cancer mortality with the stability of pancreatic, could it be that the development of new surgical techniques and disease biomarkers for early diagnoses for some cancers has received more headway than others?

Being able to visualize these differences and changes that stand out on physiology might assist us in identifying diseases with complex etymologies in the deluge of data afforded by modern toxicological and exposure science efforts. The examples of dosimetry (Figure 2), tissue-specific expression level data (Figure 3), and historical tissue-related disease phenotype data (Figure 4) all illustrate the power of annotating scientific data onto the human form.

Having PAVA as an additional visual analytics tool is vital for mapping temporal and spatial PAD, and enabling hypothesis-driven inquiry, hypothesis generation, and rationalization of complex data. We believe that these approaches could easily be extended to visualization of (a) chemical parent versus progeny (metabolite, degradation, bio-transformation) time-course studies for either exposure or effect biomarkers,(16) both endogenous and xenobiotic, (b) life-stage specific chemical disposition comparisons (e.g. adult versus child), (c) inter-species chemical disposition comparison (rat vs. human allometrically normalized), (d) other historical tissue-annotated epidemiological studies as seen in Figure 4, and model/model, model/experiment, model scenario validation and comparison. Studies are currently underway to explore and illustrate these concepts. Future PAVA developments will include consideration of various anatomies (rat, mouse, dog, chicken, fathead minnow and zebrafish) and consideration of additional organs/tissues. While the examples shown here only begin to hint at the possibilities of using visual analytics within pharmaco/toxico-kinetic studies, it is the first of its kind, to the extent of the knowledge of the authors, to an engaging, platform-independent open architecture for exploring physiologically annotated data.

Supporting Information

Information regarding PAVA v1.0 development and the server-side PAVA v1.0 application installation files can be obtained on the HEASD product/tools website in the "Models" section at http://www.epa.gov/heasd/products/products.html or under the PAVA designated page (http://www.epa.gov/heasd/products/pava/pava.html). In addition, there is product development and user forum with dialog on PAVA on the US-Environmental Protection Agency's Environmental Science Connector

(http://portal.epa.gov/ESC). Interested users may join to familiarize, comment, and learn about the PAVA development cycle and provide input for desired features in future releases by contacting the authors. Included in the supporting information are (i) a screen-capture video with dialog demonstrating how a PAVA animation may be generated (ii) animated gif rendering of Figure 1 ethanol model and (iii) Figure 4C – Male-Female differential cancer mortality time-course data animation.

Acknowledgments

The authors gratefully acknowledge Rachel Brady (Duke University Visualization Core) for providing the opportunity to present to the academic visualization community and receive insightful feedback (see http://vis.duke.edu/FridayForum/09Fall.html and http://lectopia.oit.duke.edu/ilectures/ilectures.lasso?ut=193&id=21022 **)** and Dr. Shane Peterson and Dr. Elin Ulrich (US-EPA) for thorough criticism and suggestions towards the manuscript

References

1 Wildstrom, S., A Picture is Worth a Thousand Charts: New software "visualizes" data in ways that give a much richer image than typical tools. *Business Week*, January 20 2003: http://www.businessweek.com/magazine/content/03_03/b3816045.htm.

2 Nichols, J. Rheingans, P.; Lothebach, D.; McGeachie R., Skow L., McKim J. Three-dimensional visualization of physiologically based kinetic model outputs. *Environmental Health Perspectives*, 1994. 102(11): p. 952.

3 Yoshikawa, Y., Sone, H., Yoshikawa H., Takada K. WinPBPK-A Software for Physiologically-based Pharmacokinetic Model Analysis: Application to Tissue and Organ Distribution Analysis of Tacrolimus. *Drug Metabolism and Pharmacokinetics.* 1999. 14(1): p. 22-31

4 (a) Lawton, G., Users Take a Close Look at Visual Analytics. Computer, 2009. 42(2): p. 19-22. (b) Viegas, F. B., Wattenberg, M., van Ham F., J. Kriss, and M. McKeon, Manyeyes: a site for visualization at internet scale. *IEEE Transactions on Visualization and Computer Graphics*, 2007. 13(**6**): p. 1121-1128.

5 Blancato, J. N., Power, F. W., Brown, R. N., Dary, C. C., *Exposure Related Dose Estimating Model (ERDEM) A Physiologically-Based Pharmacokinetic and Pharmacodynamic (PBPK/PD) Model for Assessing Human Exposure and Risk*. EPA/600/R-06/061, June 2006
6 Levitt. DG.; PKQuest_Java: free, interactive physiologically based pharmacokinetic software package and tutorial. 1: *BMC Res Notes*. 2009 Aug 5;2(**1**):158.
7 Levitt DG. 1: PKQuest: measurement of intestinal absorption and first pass metabolism - application to human ethanol pharmacokinetics; *BMC Clin Pharmacol*. 2002 Aug 15;2:4
8 Bourne, D.W.A. 1989. BOOMER, a simulation and modeling program for pharmacokinetic and pharmacodynamic data analysis. Computer Methods and Programs in Biomedicine 29, pp 191-195.
9 Cahill, T., I. Cousins, and D. Mackay, Development and application of a generalized physiologically based pharmacokinetic model for multiple environmental contaminants. *Environmental Toxicology and Chemistry*, 2003. 22(1): p. 26-34.
10 Smith, T., N. Petoussi-Henss, and M. Zankl, Comparison of internal radiation doses estimated by MIRD and voxel techniques for a" family" of phantoms. *European Journal of Nuclear Medicine and Molecular Imaging*, 2000. 27(**9**): p. 1387-1398.
11 Weinstein, S., Magic images via Imagemagick. *The C Users Journal*, 1993. 11(**10**): p. 93-102.
12 Nishimura M.; Naito S.; Yokoi T; Tissue-specific mRNA expression profiles of Nuclear Receptor Superfamily. *Drug Metabolism and Pharmacokinetics* **19**, 2, 135-149, 2004
13 Nishimura, M,; Hiroshi Y.; Yoshitsugu, H; Naito, S.; Satoh, T.; Tissue distribution of mRNA expression of human cytochrome p450 isoforms.*Yakugaku Zasshi* 123(**5**) 369-375 (2003)
14 Nishimura M.; Naito S.; Tissue-specific mRNA expression profiles of Human Phase 1 metabolizing enzymes except for cytochrome p450 and Phase II metabolizing Enzymes. *Drug Metabolism and Pharmacokinetics* 21(**5**) 357-374, 2006
15 US Mortality Data 1960-2005, US Mortality Volumes 1930-1959, National Center for Health Statistics, Centers for Disease Control and Prevention, 2008. adapted from http://www.cancer.org/downloads/PRO/Cancer_Statistic_2009_Slides_rev.ppt
16 Ankley, G., N. Denslow, and K. Watanabe, A computational model of the hypothalamic-pituitary-gonadal axis in male fathead minnows exposed to 17α-ethinylestradiol and 17β-estradiol. *Toxicological Sciences* 109(**2**), 180–192 (2009)