CHAD Explorer: an enhanced Consolidated Human Activity Database (CHAD)

Shi V. Liu and Thomas McCurdy
Human Exposure and Atmospheric Science Division, National Exposure Research Laboratory, ORD, USEPA, RTP, NC 27711, USA

To be submitted to the Journal of Exposure Science and Environmental Epidemiology (JESEE)

Address all correspondence to: Dr. Liu, Exposure Modeling Research Branch, National Exposure Research Laboratory, E205-02, EPA, Research Triangle Park, NC 27711. Tel: 919 541 9410. Fax: 919 541 4787. E-mail: liu.shi@epa.gov. Dr. Liu is also the contact for obtaining a CHAD Explorer user ID and password.

Abstract

A new database-driven web application called CHAD Explorer has been developed by EPA's National Exposure Research Laboratory based upon the Laboratory's web-based Consolidated Human Activity Database (CHAD). CHAD Explorer takes advantage of the Oracle Application Express™ (OAE) as an integrated development environment, and provides various routes for browsing and querying the sequential 24-h time/activity data contained in CHAD. It also provides alternative procedures to instantly download retrieved data and exchange a variety of data and information. Using an enhanced administration module of CHAD Explorer within OAE, recently acquired time/activity data have been added to the CHAD data file. It is expected that as more time/activity data become available, they also will be added into CHAD using the OEA platform and this will provide improved functionality to the exposure modeling community.

Keywords: CHAD, CHAD Explorer, Database, Exposure Modeling, Human Activity, Time use,

Introduction

This paper discusses improvements to the US Environmental Protection Agency's (EPA's) Consolidated Human Activity Database (CHAD) available at www.epa.gov/chadnet1 (McCurdy et al., 2000). CHAD was developed to support a number of "event-based" human exposure models that utilize a sequential time-series approach to estimating exposures and intake dose rates for individuals as they go through their daily lives. To facilitate that, CHAD currently contains individual diary-days of time use for almost 23,000 person-days in its online version, and another 12,000 person-days of data available upon request. As described here, these data have now been combined into a new, interactive and publically-available database (CHAD Explorer) that will be of use to the exposure and time use research communities. CHAD uses 115 location codes and 144 activity codes, arranged in a hierarchal manner. Many exposure modeling applications use some subset of these location and activity codes to assign pollutant concentrations and ventilation rates to simulated individuals (McCurdy & Graham, 2003). To set the stage for our discussion of the new CHAD Explorer, we provide a brief history of how human activity information developed in EPA over the years.

History of human activity information in EPA

The need for activity pattern data in exposure model simulations was recognized early by Duan (1982), Fugas (1976), and Ott (1982, 1983). The first data base of sequential, time-series activity information was developed in 1979 for EPA's Office of Air Quality Planning and Standards (OAQPS) by researchers from Stanford Research Institute (SRI) (Roddin, et al., 1979). SRI disaggregated daily time use data from John Robinson's book, *How Americans Use Time* (Robinson, 1977) into age, weekday/weekend, climate category, and work-shift category "cohorts." This disaggregation was bolstered using international time-use data from an influential book at the time (Szalai 1972). The SRI database essentially was based on "best engineering judgment," as there were no large-scale time use databases available to develop sequential activity patterns. In 1984, the SRI data were subsequently subdivided by PEI International, a consultant to OAQPS, into 10-minute block diaries, again using best judgment (Pope, 1986). That database was shortly superseded by "real" time use information that started to become available in the mid-1980's from EPA's Office of Research and Development (ORD). ORD sponsored two personal exposure monitoring studies in Denver CO and Washington DC that collected "real-time" location and activity diary data from study participants (Akland et al., 1985; Hartwell et al., 1984; Ott et al., 1998). Time use diary data was significantly expanded by a 1985 Electric Power Research Institute diary study of residents living in the metropolitan Cincinnati area (Johnson, 1989). In 1987-1988 the California Air Resources Board sponsored an *ex post* (recall of "yesterday's" activities) probability-based survey of the State's residents' activity patterns (Robinson et al., 1989; Wiley et al., 1991a,b). In 1991, OAQPS combined data from the Cincinnati, Denver, and Washington DC metropolitan areas into a "Multicity Activity Data Base" (Paul et al., 1991). These data were used in a number of OAQPS air exposure modeling analyses of alternative National Ambient Air Quality Standards done in the 1990's.

Starting in 1994, EPA's ORD undertook a national probability-based time use survey of U.S. residents in the 48 contiguous states, called the "National Human Activity Pattern Survey" (NHAPS). The NHAPS database is well documented and has been extensively analyzed (Klepeis et al., 1995, 2001; Tsang & Klepeis, 1996). Since NHAPS, EPA has obtained time use data from smaller diary studies collected in-house and from its field exposure monitoring work which include a Baltimore elderly study (Williams et al., 2000); 3 studies of subjects living in the Research Triangle Park NC area (Williams et al. 2003; Rea et al., 2001 and Isaacs et al., 2011), and a study in Seattle Liu et al., 2003) and from children's time use surveys conducted by the University of Michigan as part of the Child Development Supplement of its Panel Study of Income Dynamics (www.isr.umich.edu/frc/). About 34% of data in the expanded version of CHAD comes from the 2000-2008 time period, and another 49% is from the 1990's. About half of information is cross-sectional in nature, being one diary day per person, while the other half has between 2-369 days of longitudinal data per person.

The Use of CHAD

CHAD is the time use data input file for OAQPS's APEX model that is used for many NAAQS reviews (e.g., carbon monoxide, sulfur dioxide, ozone, and nitrogen dioxide); see www.epa.gov/ttn/naaqs for more information. CHAD is the basis for human activity data for OAQPS's TRIM.Expo and HAPEM exposure models (www.epa.gov/ttn/fera/data/apex; Palma, 1999). CHAD also is used in various versions of ORD's SHEDS model (e.g., Burke et al., 2001; Xue et al., 2006; Zartarian et al., 2006). The MENTOR exposure and dose computational system also uses CHAD data in its exposure module (Georgopoulos & Lioy, 2006; Georgopoulos et al., 2005, 2008). Non-EPA analysts use the data in CHAD for a number of purposes, including descriptive analyses of time use (Barraj et al, 2007; Mokhtari & Frey, 2005); development of time-averaged statistics of time use (Hollowman et al., 2004); serving as reference of average time-activity patterns of general US population in comparing with those patterns of specific local populations (Wu et al., 2010), and as input to regional exposure models (Lu, et al., 2005). A March 2011 Google Scholar™ search indicates that CHAD is cited in over 520 articles and books, so the database is relatively well known in the exposure modeling literature.

The current presentation of CHAD

The current presentation of CHAD has two formats: a Microsoft Access™ stand-alone database and a web-accessible database contained within Oracle's™ database management system (Glen et al., 1997). However, due to technical constraints associated with the original implementation of CHAD and changes made to it over time, it is difficult and time-consuming to add data to the database in its current formats. In addition, CHAD users have indicated that they would like more functionality in using it, including being able to query individual parts of it more easily.

To address these issues, a new version of CHAD, called CHAD Explorer, has been developed. CHAD Explorer combines the advantages of the relational database implemented in the original Access version of CHAD and the web accessibility afforded by the Oracle database version of CHAD. In addition, more data-mining capabilities are built into this new web-accessible relational database application. Most importantly, CHAD Explorer was built and is maintained

completely by EPA in-house by taking advantages afforded by Oracle Application Express™ (OAE) (Oracle, 2010). This more easily sustainable feature of CHAD Explorer is cost-effective and is critical for enriching CHAD with more data and enhancing CHAD Explorer with more functionality.

Materials and Methods

The Object Browser of OAE (Oracle, 2010) was used for creating CHAD Explorer tables and other viewed objects in the new database. The original CHAD data were stored in two large flat tables using Oracle's Database Management System™ (DBMS); these data were partitioned and marshaled into new tables through a combination of OAE data-exporting and importing tools. Additional new CHAD data was taken from a master file compiled by Alion, a contractor to EPA. The new web application was also developed by using OAE. The following topics were addressed in the formulation of CHAD Explorer.

Results

Restored a relational database for efficient storage of CHAD data. To reduce data redundancy and thus increase storage efficiency, the current web-accessible CHAD data residing in two large flat tables in an Oracle database, was restored as a relational database structure similar to the standalone Microsoft Access database version. CHAD Explorer's new entity-relationship diagram is presented in Figure 1. The diagram depicts 11 different entities, 7 of them oriented toward personal characteristics and accessed by the unique identification number for each person in the file. The top row includes selected health factors (see "PersonHealth," top left), housing characteristics (PersonHouse, middle), and quality assurance/quality control "flags" focused on the *person*. These flags were used in the original CHAD to highlight potential problems with an individual's data, such as too much or too little sleep, missing location or activity entries, and the number of records for each individual. The flag on number of records can be compared to the "criterion" of a compliant individual (>30 records) defined by Graham & McCurdy (2004). A flag does not prohibit the analyst from using the datum; it just highlights a potential issue that the user should be aware of.

In the second row going left-to-right, the first box depicts a "Study"-specific entity, used if someone wants to query CHAD Explored by Study Name (NHAPS, Baltimore, etc.). The next two entities are again person-oriented; "StudyPerson" is related to selected physical characteristics, such as age, gender, weight, and ethnicity, while "PersonTime" provides information on the specific dates for which the diary day data were gathered, including "day type" (weekday/Saturday/Sunday), and how many days of data are available for any single individual. The right two boxes depict the person-specific daily sequence of events undertaken. "TimeActivity" information includes a location and activity code, each event's start and stop clock times, whether or not smoking or gas stove usage occurred, and whether or not the person was breathing hard during an event. "ActivityQC" provides information on qualitly control flags for *individual data entries*.

The bottom row depicts CHAD Explorer entities focused on other characteristics of the individual's diary data. "PersonLocal" provides information on geographical location, such as

county and state name, and her or his residential ZIP code. "PersonSocial" depicts educational and job-related information, including income, but there are very few entries for this datum and the information is dated anyway. "TimeWeather" provides daily meteorological data for each diary day of data; included are average and maximum daily temperature and precipitation data. This type of information is used in the APEX and SHEDS models as "pooling variables" when cohorts are developed to better mimic human activities that change when the weather changes.

The two bottom right boxes are "embedded cross-references" to link location and activity descriptors to each specific event in a person's diary. The exact meaning of each code can be looked up in the code description table listed under the Help menu.

Developing alternative browsing approaches. CHAD Explorer contains a graphical user interface (GUI) that facilitates the intuitive exploration of data. It operates in a tab-list fashion that supports two ways of browsing the data: in a flat view version that depicts the data in a linear manner. See Figure 2. The figure shows the "flat view list" of available data tables (study, weather, codes used (Figure 2A) and the type of information available when "Study" is selected. Figure 2B, on the other hand, depicts the "Drill Down" data screen and the type of information available when Baltimore is chosen as the Study ID.

The GUI allows the query of different aspects of CHAD data according to the criteria set up by the user during a "Query" (the Tab next to "Browse"). See Figure 3. Because CHAD Explorer data are stored in small relational tables, a query result can be quickly obtained by searching a small table containing only relevant data rather than searching through a big table containing all of the CHAD data, as was the case with the original version of CHAD. A "view object" Tab that pops up when Query is chosen, depicts all of the different aspects of CHAD data in the entities discussed earlier. These include the person's personal, home, and social data, as well as meteorological and time use information. When Query is selected, the "Selection" box pops up. If the default "ALL' is chosen the information shown in the top right box is provided. If any criterion is selected, then the query result will be reflected in the bottom reporting table.

Facilitate data exchange with the time use and exposure research community. CHAD Explorer is designed to foster data exchange with time use researchers. Thus, a Download function is provided under each table showing the browsed or queried results which can be downloaded into user's own computer as a CSV (comma-separated value) file. In addition, a centralized "File Transfer" module is provided at the top level GUI of the CHAD Explorer. This "File Transfer" module includes a "Download" and an "Upload" function. The download function allows obtaining any kind of information deemed as useful for human activity study and thus has been archived into a specific database table in CHAD Explorer. Using the same file transfer module, but the "Upload" function, users can upload to CHAD Explorer their information in a variety of file formats. Figure 4 depicts the pop up for uploading data, including a file name and a general description. The file formats allowed for uploading include CSV, ASCII, PDF, Excel™ spreadsheets, SAS™ files, SPSS™ files, and various image files such as GIF, IPG, and NGP. All data/files submitted to CHAD Explorer will undergo quality control analyses by the Database Manager and are also subject to check on adherence to relevant policy and regulation prior to new information being placed on EPA's public server for dissemination.

Get feedback and provide help on a regular basis. CHAD Explorer provides a help and feedback module (Figure 5). For example, abbreviations used, a glossary, and frequently asked questions (FAQs) are provided on a toolbar. There are also lists of table and column names, and various codes used for classification of data. The contents of this information will be built up dynamically over time, but the interface for accessing the content will remain stable for familiar usage.

Provide an administration module for backend database management. CHAD Explorer also provides a powerful backend database management module that allows the system's administrator to monitor and control all aspects of the database. It should be pointed out that because most of the GUI components (e.g., the various icons and images) are based on files stored in the database tables themselves, simply changing the file content in any database table can achieve some face-lifting of CHAD Explorer easily. Similarly, updating descriptions of the tables, columns, and codes can be made without reprogramming.

Enrich CHAD with more data. With the above data management capability built into CHAD Explorer, adding more human activity data into CHAD is much easier than before. Of course, the data to be added must first be quality checked and brought to the format consistent with the database structure of CHAD. The total number of person-days of activity in CHAD Explorer is 34,773, see Table 1. The studies added to CHAD via CHAD Explorer are noted with an asterisk. CHAD Explorer was used to add another 12,057 person-day data to the original-CHAD, an increase of 34.7% more daily data. Much of the new diary data has more than one day of data per person; because of that the number of individuals with 1 or more diary days only increased from 17,738 to 21,112, a 16.0% increase. The total number of records increased from 868,819 to 1,488,243, a 41.7% increase.

Using the browsing and querying functions of CHAD Explorer, some quality assurance and quality control evaluations were performed to verify that the right numbers of records are returned by these data-mining tools.

Discussion

The advancement of information technology since CHAD was first developed in 2000 has provided significant possibilities for developing CHAD Explorer. For example, the OAE from Oracle (Oracle 2010), with its new integrated development environment (IDE), greatly facilitated the construction of CHAD Explorer. This allowed EPA staff to develop the expanded database web application themselves rather than relying on external contractual support, thus saving EPA money and allowing the Agency to better control the product and future expansion of it. In addition, the new database web application more easily allows versatile exploration of time use data in CHAD and interfacing with the system to download or upload new data. An offshoot of the project is more efficient storage of time/activity data.

Using the capability of OAE and CHAD Explorer, some additional human activity data was added into CHAD. The addition of 8 new small time use studies increased the number of subjects contained in the database by 16%, increased by 35% the number of person-days of diary data, and increased by 42% the number of respective type of records in CHAD. The differences

among these percentages reflects the new emphasis on longitudinal time/activity data, which is very critical for exposure assessments due to the need to address both intra- and inter-individual variability in a person's time use (Glen et al., 2007). Ignoring intra-individual variability systematically underestimates the "high-end" of the population exposure distribution because it is so dependent upon repetitive activities of exposed people (Law et al., 1997; McCurdy 1997).

In addition, we want to make it easy to incorporate more human activity data originated from new, non-EPA research into CHAD. Some of these data already exist in repositories that are not directly compatible with CHAD and thus will require some changes in their data structure and coding scheme to be integrated into CHAD Explorer. The American Time Use Survey (ATUS) may be an example of such a database (www.bls.gov/tus/), but see George & McCurdy (2010) for a discussion of issues associated with using it for exposure modeling.

We have tried to develop a flexible structure in CHAD Explorer to facilitate the acquisition of data that are not directly compatible with the current CHAD, and make it more flexible and easier to use. This increase in flexibility might even point a way for incorporating into CHAD other human activity data from other countries such as the Harmonized European Time Use Study (HETUS) and the mostly European Multinational Time Use Study (Fisher & Gershuny, 2008; Fisher & Robinson, 2009, 2010; Gershuny, 2004; Leech et al., 2002). We think that our work is a step in that direction and we will try to maintain the development of CHAD-related research tools in a sustainable manner. CHAD Explorer can be accessed at the same URL as CHAD: www.epa.gov/chadnet1/.

Acknowledgements

Staff of Alion Scientific, a contractor to EPA, provided a copy of the CHAD "Masterfile" of SAS-based data that is the primary source for CHAD Explorer tables. We also thank Drs. Jianping Xue and Kristin Isaacs for a constructive internal review of a draft of this manuscript. Comments from the Journal's anonymous reviewers improved the clarity and content of the paper. We thank them for their efforts.

Disclaimer

This paper has been reviewed in accordance with EPA's internal procedures, and has been approved for publication. Mention of trade names or other commercial products does not constitute endorsement of their use by EPA.

Conflict of interest

The authors declare no conflict of interest. Staff of the U.S. Environmental Protection Agency developed CHAD Explorer using taxpayer-provided funds from the Agency's general research budget.

References

Akland G.G., Hartwell T.D., Johnson T.R., and Whitmore R.W. Measuring human exposure to carbon monoxide in Washington,DC, and Denver, Colorado during the winter of 1982-1983. *Environ Sci Technol* 1985: 19: 911-918.

Barraj, L.M., Tsuji, J.S., and Scrafford, C.G. The SHEDS-Wood Model: incorporation of observational data to estimate exposure to arsenic for children playing on CCA-treated wood structures. *Enviro. Health Persp* 2007: 115: 781-786.

Burke J.M., Zufall M.J., and Ozkaynak H. A population exposure model for particulate matter: case study results for PM(2.5) in Philadelphia, PA. *J Expos Anal Environ Epidem* 2001: 11: 470-489.

Duan, N. *Models for Human Exposure to Air Pollution.* Washington DC: Rand Corporation, 1982.

Fisher, K. And Gershuny, J. *Diaries from children and young people: supplement of the Multinational Time Use Study.* Oxford: Centre for Time Use Research, University of Oxford, 2008.

Fisher, K. and Robinson, J. Average weekly time spent in 30 basic activities across 17 countries. *Soc Indic Res* 2009: 93: 249-254.

Fisher, K. and Robinson, J. Daily routines in 22 countries: dairy evidence of average daily time spent in thirty activities. Paper 2010-01; Centre for Time Use Research, University of Oxford, United Kingdom, 2010.

Fugas, M. *Proceedings of the International Conference on Environmental Sensing and Assessment.* New York: Institute of Electrical and Electronic Engineers, 1976.

George B.J., and McCurdy T. Investigating the American Time Use Survey from an exposure modeling perspective. *J Expos Sci Environ Epidem* 2011: 21: 92-105.

Georgopoulos P.G., and Lioy P.J. From a theoretical framework of human exposure and dose assessment to computational system implementation: the Modeling ENvironment for TOtal Risk Studies (MENTOR). *J Toxicol Environ Health B Crit Rev* 2006: 9: 457-483.

Georgopoulos, P.G., Wang, S.-W., Vyas, VM.,Sun, Q., Burke, J., Vedantham, R., McCurdy, T., and Ozkaynak, H. A source-to-dose assessment of population exposures to fine PM and ozone in Philadelphia, PA, during a summer 1999 episode. *J Expos Anal Environ Epidem* 2005: 15: 439-457.

Georgopoulos P.G., Wang S.W., Yang Y.C., Xue J., Zartarian V.G., McCurdy T., and Ozkaynak H. Biologically based modeling of multimedia, multipathway, multiroute population exposures to arsenic. *J Expo Sci Environ Epidem* 2008: 18(5): 462-476.

Gershuny, J. MTUS progress report: the Multinational Time Use Study. *Inter J Time Use Res* 2004: 1: 101-107.

Glen, G., Smith, L., Isaacs, K., McCurdy, T., and Langstaff, J. A new method of longitudinal diary assembly for human exposure modeling. *J Expos Sci Environ Epidem* 2008: 18: 299-311.

Glen G., Lakkadi Y., Tippett J.A., and del Valle-Torres M. *Development of NERL/CHAD: The National Exposure Research Laboratory's Consolidated Human Activity Database.* Research Triangle Park, NC: ManTech Environmental Technology,1997.

Graham S.E. and McCurdy T. Developing meaningful cohorts for human exposure models. *J Expos Anal Environ Epidem* 2004: 14: 23-43.

Hartwell, T.D., Clayton, C.A., Michie, R.M., et al. *Study of Carbon Monoxide Exposure of Residents of Washington DC and Denver, Colorado.* Research Triangle Park NC: Research Triangle Institute, 1984.

Holloman, C.H., Bortnick, S.M., Morara, M, Strauss, W.J., and Calder, C.A. A Bayesian hierarchial approach for relating $PM_{2.5}$ exposure to cardiovascular mortality in North Carolina." *Environ. Health Persp* 2004: 112: 1282-1288.

Isaacs, K., McCurdy, T., Glen, G., Nyswander, M., Errickson, A., Forbes, S., Graham, S., McCurdy, L., Smith, L., Tulve, N., and Vallero, D. Statistical properties of longitudinal time/activity data for use in human exposure models. *J. Exp. Anal. Environ. Epidem.* (submitted February 2011).

Johnson, T. *A Study of Human Activity Patterns in Cincinnati, Ohio.* Palo Alto CA: Electric Power Research Institute, 1989.

Klepeis N.E., Nelson W.C., Ott W.R., Robinson, J.P., Tsang, A.M., Switzer, P., Behar, J.V., Hern, S.C., and Engelmann, W.H. The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. *J Expo Anal Environ Epidem* 2001: 11: 231-252.

Klepeis, N., Tang, A., and Behar, J.V. *Analysis of the National Human Activity Pattern Survey (NHAPS): Responses from a Standpoint of Exposure Assessment.* Las Vegas: US Environmental Protection Agency, 1995.

Law, P.L., Lioy, P.J., Zelenka, M.P., Huber, A.H., and McCurdy, T.R. Evaluation of a probabilistic exposure model applied to carbon monoxide (PNEM/CO) using Denver personal exposure monitoring data. *J Air Waste Manage Assoc* 1997: 47: 491-500.

Leech J.A., Nelson W.C., Burnett R.T., Aaron S., and Raizenne M.E. It's about time: a comparison of Canadian and American time-activity patterns. *J Expo Anal Environ Epidem* 2002: 12: 427-432.

Liu, L.-J. S., Box, M., Kalman, D., Kaufman, J., Koenig, J., Larson, T., Lumley, T., Sheppard, L., and Wallace, L. (2003). "Exposure assessment of particulate matter for susceptible populations in Seattle." *Environ. Health Persp.* 111: 909-918.

Lu, R., Wu, J., Turco, R.P., Winer, A.M., Atkinson, R., Arey, J., Paulson, S.E., Lurmann, F.W., Miguel, A.H., and Eigurer-Fernandez, A. Naphthalene distributions and human exposure in Southern California. *Atmos Environ* 2005: 39: 489-507.

Mokhtari, A. and Frey, C. (2005). Sensitivity analysis of a two-dimensional probabilistic risk assessment model using analysis of variance. *Risk Analysis* 2005: 25: 1511-1525.

McCurdy, T. Modeling the dose profile in human exposure assessments: ozone as an example. *Rev Tox: In Vivo Tox Risk Assess* 1997: 1: 3-23.

McCurdy T., Glen G., Smith L., and Lakkadi Y. The national exposure research laboratory's consolidated human activity database. *J Expo Anal Environ Epidem* 2000: 10: 566-578.

McCurdy T., and Graham S.E. Using human activity data in exposure models: analysis of discriminating factors. *J Expo Anal Environ Epidem* 2003: 13: 294-317.

Oracle. Oracle Application Express. *http://apexoraclecom/i/indexhtml* 2010.

Ott, W.R. Concepts of human exposure to air pollution. *Environ Inter* 1982 7: 179-196.

Ott W.R. Exposure estimates based on computer generated activity patterns. *J Toxicol Clin Toxicol* 1983: 21: 97-128.

Ott W.R., Thomas J., Mage D., and Wallace L. Validation of the simulation of human activity and pollutant exposure (SHAPE) model using paired days from the Denver, CO, carbon monoxide field study. *Atmos Environ* 1998: 22: 2101-2113.

Palma, T., Vasu, A.B., and Hetes, R.G. The Total Risk Integrated Methodology (TRIM). EM 1999 (March): 30-34.

Paul, R.A., Yoder, J.M., Johnson, T., et al. *Documentation for the Multicity Activity Data Base.* Durham NC: International Technology Air Quality Services, 1991.

Pope, A. *Development of Activity Patterns for Population Exposure to Ozone.* Durham NC: PEI Associates, 1986.

Rea, A.W., Zufall, M.J., Williams, R.W., Sheldon, L., and Howard-Reed, C. The influence of human activity patterns on personal PM exposure: A comparative analysis of filter-based and continuous particle measurements. *J. Air Waste Mange. Assoc.* 2001: 51: 1271-1279.

Robinson, J. *How Americans Use Time.* New York: Praeger Publishers, 1977.

Robinson, J., Wiley, J.A., Piazza, Y., et al. *Activity Patterns of California Residents and Their Implications for Potential Exposure to Pollution.* Sacramento CA: California Air Resources Board, 1989.

Roddin, M.F., Ellis, H.T., and Siddiqee, W.M. *Background Data for Human Activity Pattern: Vols. I and II.* Palo Alto CA: Stanford Research Institute, 1979.

Szalai A. *The Use of Time: Daily Activities of Urban and Surburban Populations in Twelve Countries.* The Hague: Mouton, 1972.

Tsang, A.M. and Klepeis, N.E. *Descriptive Statistics Tables from a Detailed Analysis of the National Human Activity Pattern Survey (NHAPS) Data.* Las Vegas: US Environmental Protection Agency, 1996.

Williams, R., Suggs, J., Zweidinger, R., Evans, G., Creason, J., Kwok, R., Rodes, C., Lawless, P., and Sheldon, L. The 1988 Baltimore particulate matter epidemiology-exposure study: Part1-comparison of ambient, residential outdoor, indoor and apartment particulate matter monitoring. *J. Exp. Anal. Environ. Epidem.* 2000: 10: 518-532.

Williams, R., Suggs, J., Zweidinger, R., Evans, G., Creason, J., Kwok, R., Rodes, C., Lawless, P., and Sheldon, L. The 1988 Baltimore particulate matter epidemiology-exposure study: Part1-comparison of ambient, residential outdoor, indoor and apartment particulate matter monitoring. *J. Exp. Anal. Environ. Epidem.* 2000: 10: 518-532.

Wiley, J.A., Robinson, J.P., Cheng, Y.-T., et al. *Study of Children's Activity Pattern.* Berkeley CA: University of California, 1991a.

Wiley, J.A., Robinson, J.P., Piazza, T., et al. *Activity Patterns of California Residents.* Berkeley: University of California, 1991b.

Wu, X.-M., Fan, Z.-H., Ohman-Strickland, P., Time-location patterns of a population living in an air pollution hotspot. J. Environ. Public Health 2010. 2010 Article ID625461.

Xue J., Zartarian V.G., Ozkaynak H., Dang W., Glen G., Smith L., and Stallings C. A probabilistic arsenic exposure assessment for children who contact chromated copper arsenate (CCA)-treated playsets and decks, Part 2: Sensitivity and uncertainty analyses. *Risk Anal* 2006: 26: 533-541.

Zartarian V.G., Xue J., Ozkaynak H., Dang W., Glen G., Smith L., and Stallings C. A probabilistic arsenic exposure assessment for children who contact CCA-treated playsets and decks, Part 1: Model methodology, variability results, and model evaluation. *Risk Anal* 2006: 26: 515-531.

Table 1. Summary of the CHAD database

Study Name	Year*	No. of Diaries	Number of Days Of Data per Person: Range	Median	Sponsor
Denver MSA	1983	805	1	1	EPA
Washington DC MSA	1983	699	1	1	EPA
Cincinnati MSA	1986	2,614	1 - 3	3	EPRI
California - adolescents	1988	183	1	1	CARB
California - adults	1988	1,579	1	1	CARB
Los Angeles - elementary	1989	51	3	3	API
Los Angeles - high school	1990	43	2 - 3	3	API
California - children	1990	1,200	1	1	CARB
Valdez, Alaska	1991	397	1	1	Oil Companies
NHAPS - A	1994	4,723	1	1	EPA
NHAPS - B	1994	4,663	1	1	EPA
PSID (CDS) I	1997	5,616	1 - 2	2	NICHHD
Baltimore Elderly	1998	391	1 - 24	14	EPA
EPA #1*	2000	367	367	367	EPA
RTP Unhealthy	2001	1,000	8 - 33	32	EPA
Seattle MSA*	2002	1,693	5 - 10	10	EPA
EPA #2*	2002	197	197	197	EPA
PSID (CDS) II	2003	4,782	1 - 2	2	NICHHD
RTI Averting Behavior*	2003	2,907	1 - 6	4	EPA
Internal EPA*	2007	432	35 - 69	54	EPA
EPA #1*	2007	369	369	369	EPA
Mother & Child*	2008	62	31	31	EPA
Totals		34,773			

Notes & Abbreviations

* The last year of a multi-year study is used.

Number (of days)

API = American Petroleum Institute

CARB = California Air Resources Board

CDS = Child Development Supplement

EPA = Environmental Protection Agency

EPRI = Electric Power Research Institute

MSA = Metropolitan Statistical Area

NHAPS = National Human Activity Pattern Survey (A=air version; B=water version)

NICHHD = National Institute of Child Health and Human Development

PSID = Population Study of Income Dynamics

RTI = Research Triangle Institute

RTP = Research Triangle Park

Figure legends

Figure 1. Entity relationship diagram (ERD) for CHAD Explorer.

Figure 2. Browsing CHAD in different ways. (A) Flat-view; (B) Drilling down.

Figure 3. Querying CHAD from different perspectives.

Figure 4. Data exchange options.

Figure 5. Online help information to facilitate usage of CHAD Explorer.

CHAD Explorer Figures

Figure 1.

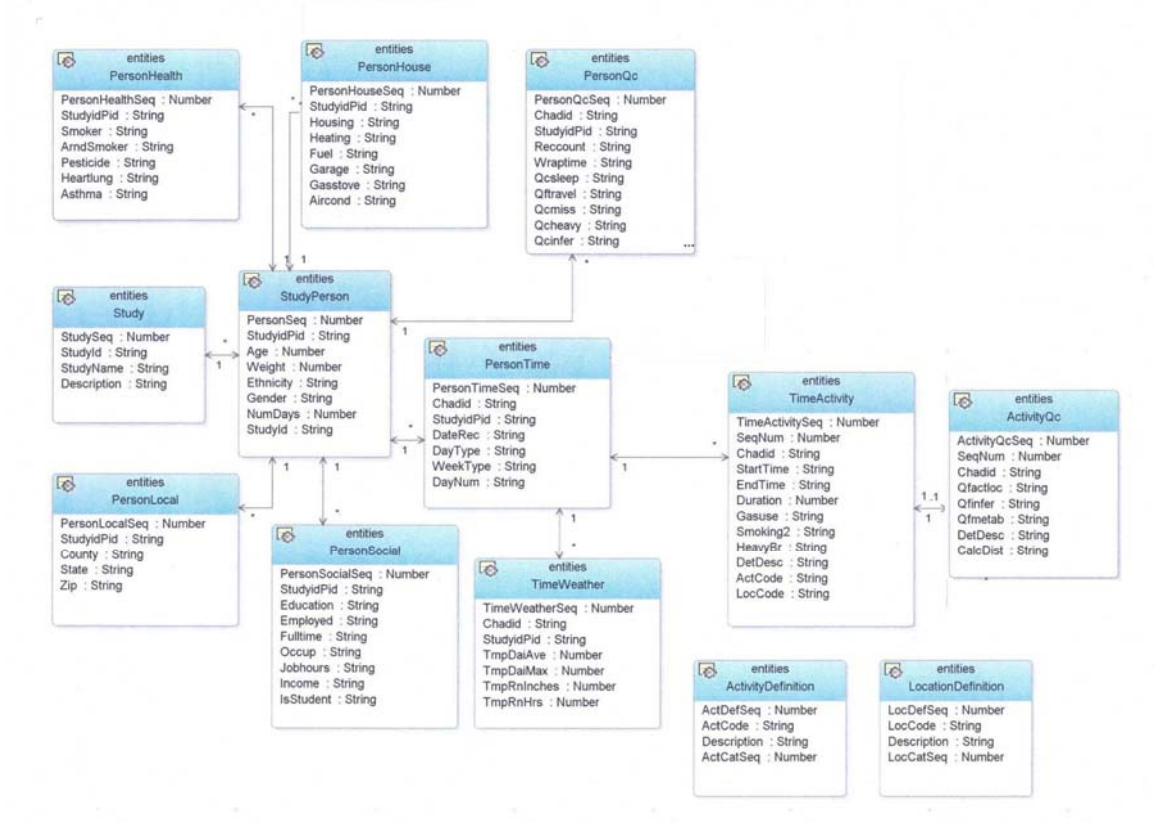
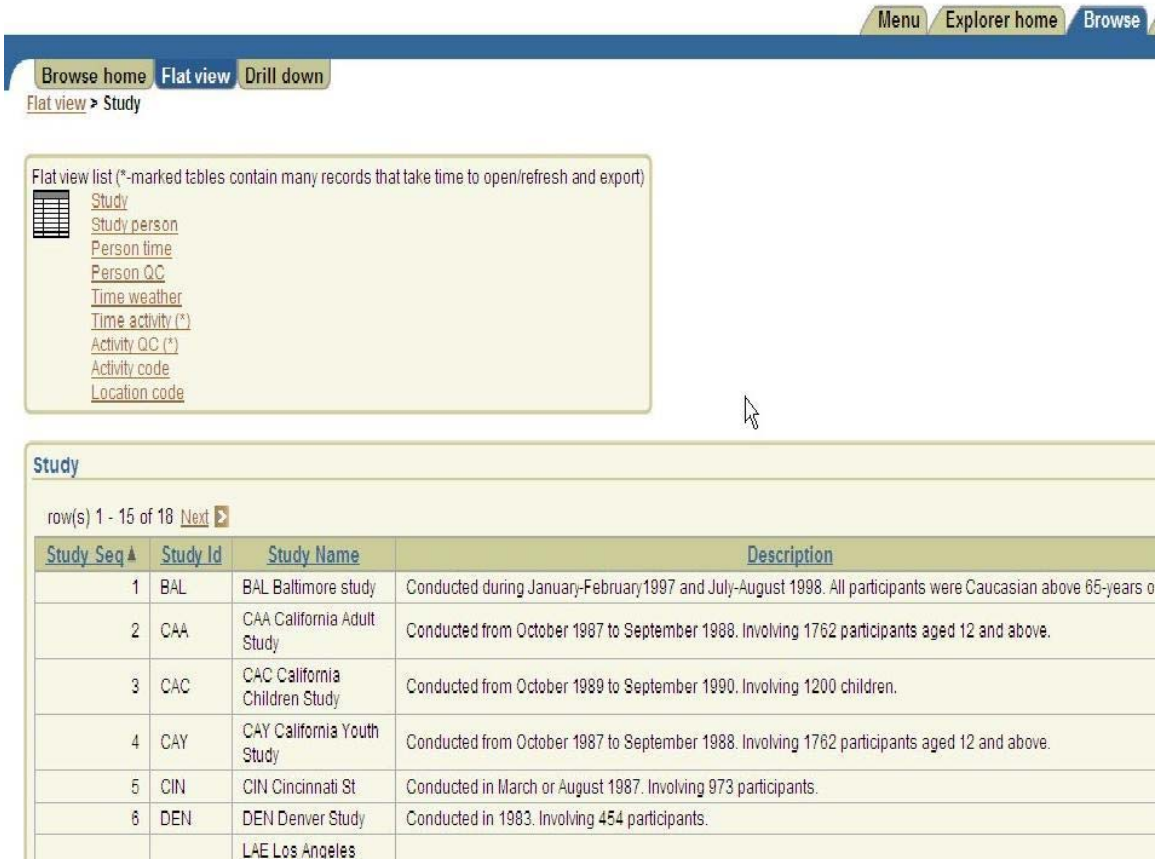


Figure 2A.



Study Seq	Study Id	Study Name	Description
1	BAL	BAL Baltimore study	Conducted during January-February1997 and July-August 1998. All participants were Caucasian above 65-years o
2	CAA	CAA California Adult Study	Conducted from October 1987 to September 1988. Involving 1762 participants aged 12 and above.
3	CAC	CAC California Children Study	Conducted from October 1989 to September 1990. Involving 1200 children.
4	CAY	CAY California Youth Study	Conducted from October 1987 to September 1988. Involving 1762 participants aged 12 and above.
5	CIN	CIN Cincinnati St	Conducted in March or August 1987. Involving 973 participants.
6	DEN	DEN Denver Study	Conducted in 1983. Involving 454 participants.
		LAE Los Angeles	

Figure 2B.

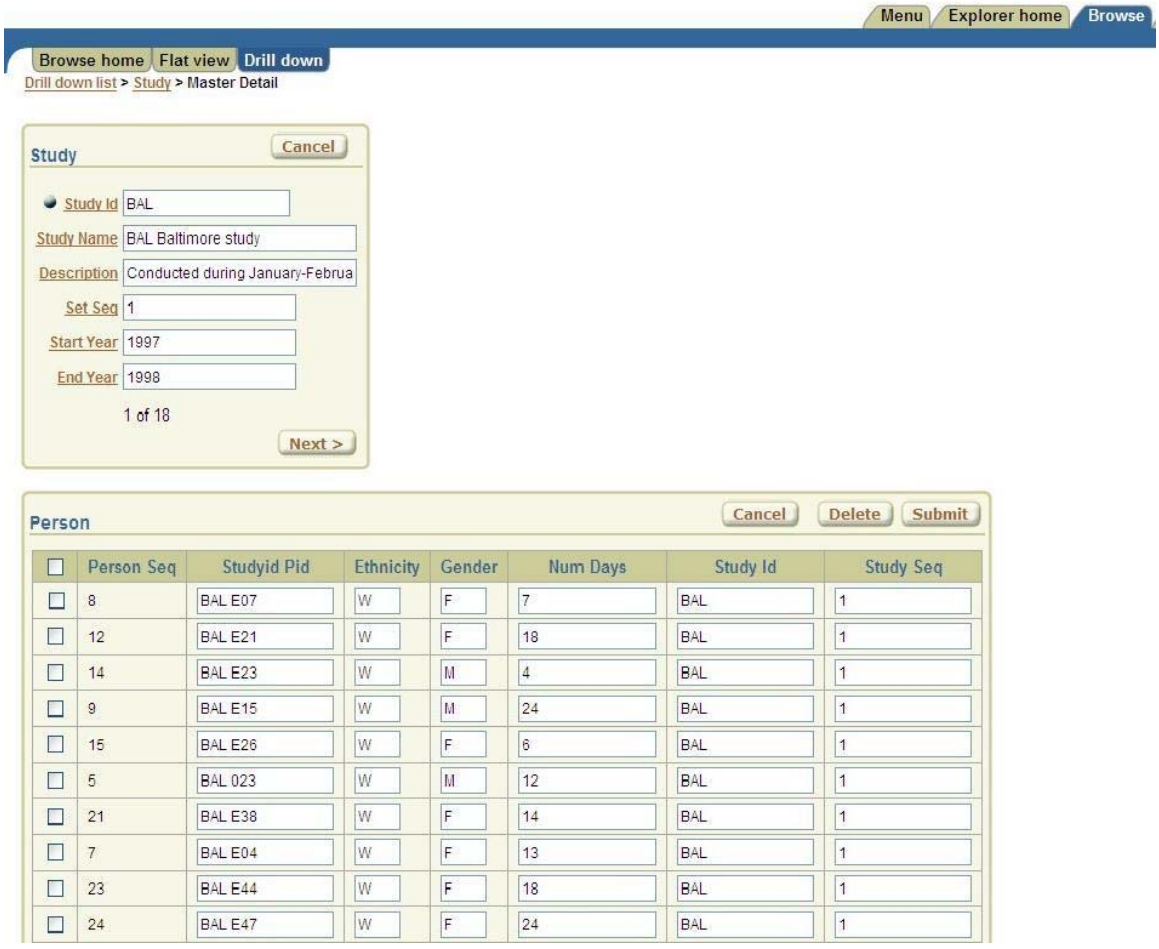

Menu
Explorer home
Browse
Browse home
Flat view
Drill down
Drill down list > Study > Master Detail
Study
Cancel
Study Id BAL
Study Name BAL Baltimore study
Description Conducted during January-Februa
Set Seq 1
Start Year 1997
End Year 1998
1 of 18
Next >
Person
Cancel
Delete
Submit
Person Seq
Studyid Pid
Ethnicity
Gender
Num Days
Study Id
Study Seq
8 BAL E07 W F 7 BAL 1
12 BAL E21 W F 18 BAL 1
14 BAL E23 W M 4 BAL 1
9 BAL E15 W M 24 BAL 1
15 BAL E26 W F 6 BAL 1
5 BAL 023 W M 12 BAL 1
21 BAL E38 W F 14 BAL 1
7 BAL E04 W F 13 BAL 1
23 BAL E44 W F 18 BAL 1
24 BAL E47 W F 24 BAL 1

Figure 3.

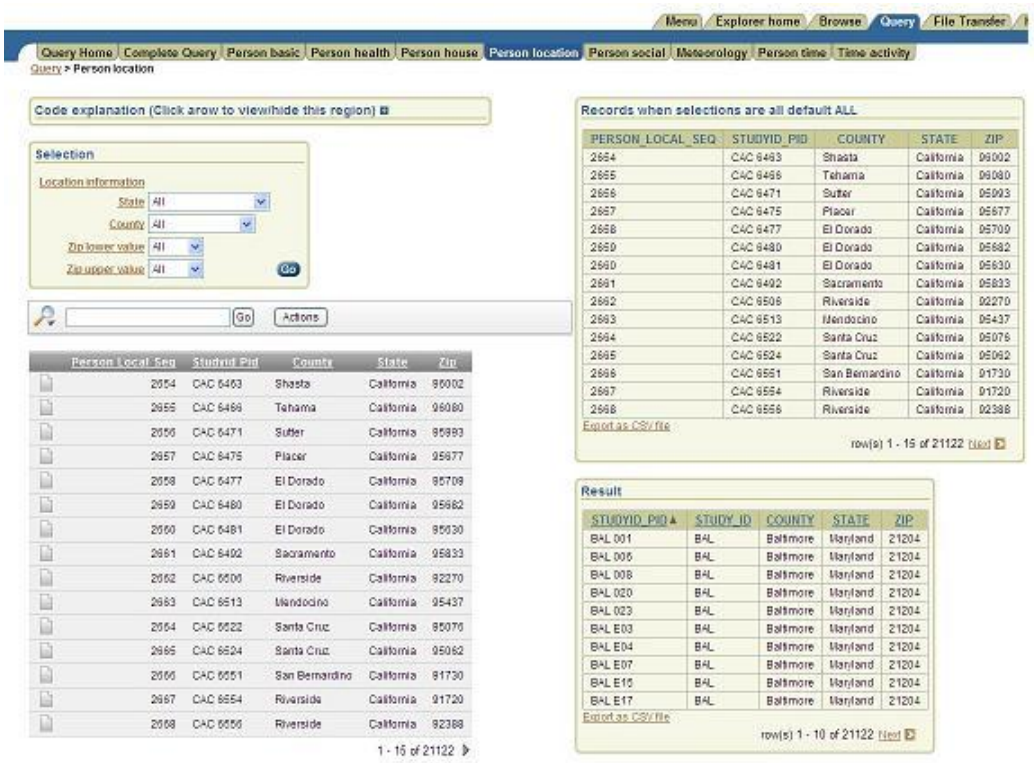
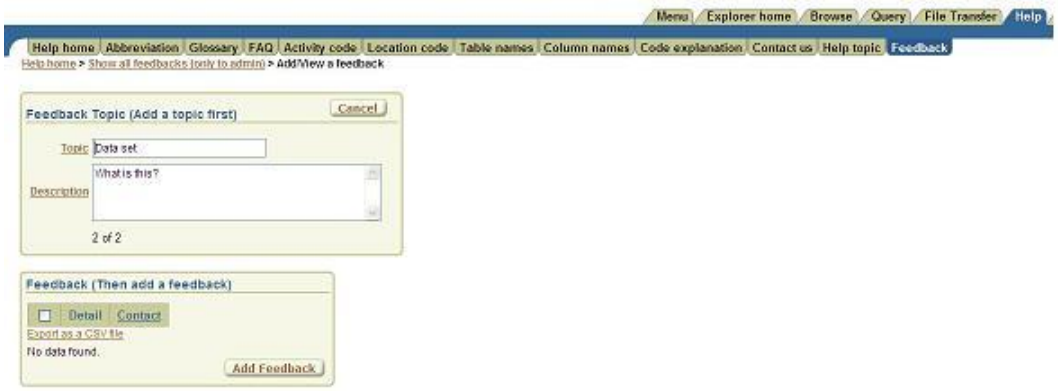

Menu
Explorer home
Browse
Query
File Transfer
Query Home
Complete Query
Person basic
Person health
Person house
Person location
Person social
Meteorology
Person time
Time activity
Query > Person location
Code explanation (Click arow to view/hide this region)
Selection
Location information
State All
County All
Zip lower value All
Zip upper value All
Go
Go
Actions
Person Local Seq
Study id Pid
County
State
Zip
2654 CAC 6463 Shasta California 96002
2655 CAC 6466 Tehama California 96080
2656 CAC 6471 Sutter California 95993
2657 CAC 6475 Placer California 95677
2658 CAC 6477 El Dorado California 95709
2659 CAC 6480 El Dorado California 95682
2660 CAC 6481 El Dorado California 95630
2661 CAC 6492 Sacramento California 95833
2662 CAC 6506 Riverside California 92270
2663 CAC 6513 Mendocino California 95437
2664 CAC 6522 Santa Cruz California 95076
2665 CAC 6524 Santa Cruz California 95062
2666 CAC 6551 San Bernardino California 91730
2667 CAC 6554 Riverside California 91720
2668 CAC 6556 Riverside California 92388
1 - 15 of 21122
Records when selections are all default ALL
PERSON_LOCAL_SEQ STUDYID_PID COUNTY STATE ZIP
2654 CAC 6463 Shasta California 96002
2655 CAC 6466 Tehama California 96080
2656 CAC 6471 Sutter California 95993
2657 CAC 6475 Placer California 95677
2658 CAC 6477 El Dorado California 95709
2659 CAC 6480 El Dorado California 95682
2660 CAC 6481 El Dorado California 95630
2661 CAC 6492 Sacramento California 95833
2662 CAC 6506 Riverside California 92270
2663 CAC 6513 Mendocino California 95437
2664 CAC 6522 Santa Cruz California 95076
2665 CAC 6524 Santa Cruz California 95062
2666 CAC 6551 San Bernardino California 91730
2667 CAC 6554 Riverside California 91720
2668 CAC 6556 Riverside California 92388
Export as CSV file
row(s) 1 - 15 of 21122 Next
Result
STUDYID_PID STUDY_ID COUNTY STATE ZIP
BAL 001 BAL Baltimore Maryland 21204
BAL 005 BAL Baltimore Maryland 21204
BAL 008 BAL Baltimore Maryland 21204
BAL 020 BAL Baltimore Maryland 21204
BAL 023 BAL Baltimore Maryland 21204
BAL E03 BAL Baltimore Maryland 21204
BAL E04 BAL Baltimore Maryland 21204
BAL E07 BAL Baltimore Maryland 21204
BAL E15 BAL Baltimore Maryland 21204
BAL E17 BAL Baltimore Maryland 21204
Export as CSV file
row(s) 1 - 10 of 21122 Next

Figure 4.


Menu
Explorer home
Browse
Query
File Transfer
File home
Download
Upload
File Report
File Upload
File Name (Up to 255 characters)
Browse...
Description (up to 4000 characters)
Submit

Figure 5.


Menu
Explorer home
Browse
Query
File Transfer
Help
Help home
Abbreviation
Glossary
FAQ
Activity code
Location code
Table names
Column names
Code explanation
Contact us
Help topic
Feedback
Help home > Show all feedbacks (only to admin) > Add/View a feedback
Feedback Topic (Add a topic first)
Cancel
Topic
Data set
Description
What is this?
2 of 2
Feedback (Then add a feedback)
Detail
Contact
Export as a CSV file
No data found.
Add Feedback