



PhilaSUG newsletter



Fall 2018 Meeting Announcement

PhilaSUG Fall 2018 Meeting Tuesday, October 30, 2018

The Philadelphia Area SAS Users Group Fall Meeting will be on Tuesday, October 30, 2018 at 1:00 PM, and will be hosted by Janssen Research & Development, 1400 McKean Rd, Spring House (Ambler), PA 19477. A map, detailed driving, and parking instructions are available later in the newsletter.

Registration will begin at 12:15 PM. The meeting will commence at 1 PM. Membership for the year is \$30. There are no other fees for attending PhilaSUG meetings. This year we will be accepting cash, check (made out to "PhilaSUG"), or credit/debit card (this charge will show up on your credit/debit card statement as paid to "Michael Davis / PhilaSUG"). If you are an employee of Janssen, fees for this meeting will be waived, but please complete the meeting sign-up form so that we can expedite entrance and to help us obtain a better food count. Receipts will be available at registration time. Driver's license will be required at lobby check-in. Also, please arrive before 1 PM as guests must be escorted.

Abstracts and bios can be found later in the newsletter.



Important – In order to expedite entrance and to help us obtain a better food count we are asking all attendees to please complete the required [Meeting Attendee Sign-up Form](#) by **Oct. 28th**.

Agenda

We thank the SAS Institute for providing lunch and break refreshments

12:15-1:00	Registration and lunch
1:00-1:10	Opening Remarks
1:10-2:00	Lex Jansen - The Present and Future of Define-XML
2:00 -2:30	Kevin Lee - Standards Driven Oncology
2:30-3:00	Break, posters, and refreshments
3:00-3:40	Amos Odeleye - Developing Credit Risk Score
3:40-4:10	Alex Chaplin - Proc Sort Revisited
4:10-4:40	Mark Keintz - History Carried Forward, Mixing Time Series of Differing Frequencies
4:40-5:00	Open Forum, Raffle Prizes and closing Remarks

Posters

Using R to Query Dataset From SAS Server and Generate Forest Plots - Xiang (Jei) Li

Using SAS and Excel to Create Clear, Informative and Realistic Mock Layouts in Word - Ken Maahs

Automation of Clinical Trial Result Posting to ClinicalTrials.gov and EudraCT - Sherry Meeh

Clinical Development Standards for FDA Bioresearch Monitoring (BIMO) Submissions - Denis Michel

The presenters and the PhilaSUG Executive Committee will adjourn for dinner at a nearby restaurant when the meeting concludes. You are invited to join us. The location will be announced at the meeting.

About Our Host



Janssen Research & Development provides research and development support for the pharmaceutical business units of parent company Johnson & Johnson (J&J). Janssen Research & Development, LLC discovers, develops, and delivers medicines and solutions for patients worldwide. It offers its solutions to address medical needs in therapeutic areas, such as neuroscience, oncology, immunology, infectious diseases and vaccines, and cardiovascular and metabolism.

With precision medicine, they are working on transforming the study of disease by investing in best-in-class research capabilities in genomics, biotherapeutics, biomarkers, molecular and companion diagnostics, the microbiome, disease interception, prevention and vaccine platforms.

Directions to Our Host Site

Directions to Spring House Site: 1400 McKean Rd Spring House, PA 19477

*GPS Address: 1400 McKean Rd, Ambler, PA 19002

FROM POINTS WEST

(Malvern, Harrisburg)

East on I-76/276 (PA Turnpike) to Exit 339 Fort Washington. Go North on Route 309 (first right after going through tollgate). Follow 309 North approximately 3 miles to Norristown Road/Spring House exit. Go left onto Norristown Road for $\frac{3}{4}$ mile to McKean Road. Turn right on McKean Road and drive approximately $\frac{1}{2}$ mile to the J&J Spring House Campus entrance on your right side. *

FROM POINTS NORTHEAST

(New Jersey, New York)

New Jersey Turnpike South to Exit 6 (PA Turnpike connector). West on I-276 (PA Turnpike) to Exit 339 Fort Washington. Go North on Route 309 (first right after going through tollgate). Follow 309 North approximately 3 miles to Norristown Road/Spring House exit. Go left onto Norristown Road for $\frac{3}{4}$ mile to McKean Road. Turn right on McKean Road and drive approximately $\frac{1}{2}$ mile to the J&J Spring House Campus entrance on your right side. *

FROM POINTS NORTHWEST

(Wilkes-Barre, Scranton, Allentown)

South to Northeast Extension of PA Turnpike to intersection with PA Turnpike. Follow PA Turnpike East to Exit 339 Fort Washington. Go North on Route 309 (first right after going through tollgate). Follow 309 North approximately 3 miles to Norristown Road/Spring House exit. Go left onto Norristown Road for $\frac{3}{4}$ mile to McKean Road. Turn right on McKean Road and drive approximately $\frac{1}{2}$ mile to the J&J Spring House Campus entrance on your right side. *

FROM PHILADELPHIA

(Center City)

Route I-76 West (Schuylkill Expressway) to intersection with Exit 326 Valley Forge of PA Turnpike. Take PA Turnpike East to Exit 339 Fort Washington. Go North on Route 309 (first right after going through tollgate). Follow 309 North approximately 3 miles to Norristown Road/Spring House exit. Go left onto Norristown Road for $\frac{3}{4}$ mile to McKean Road. Turn right on McKean Road and drive approximately $\frac{1}{2}$ mile to the J&J Spring House Campus entrance on your right side. *

FROM ROUTE 309 NORTH

(Quakertown, Coopersburg)

Take Route 309 South past Montgomeryville Mall. Continue South to Welsh Road, Route 63. Cross Welsh Road and take the jug handle to Welsh Road across Route 309. Continue on Welsh Road to the second traffic light, McKean Road. Turn right onto McKean Road and travel approximately $\frac{1}{2}$ mile to the J&J Spring House Campus entrance on your left. *

FROM POINTS SOUTH

(Baltimore, Washington)

North on I-95 to PA. Take Exit 7 for I-476 North towards Plymouth Meeting. Take Exit onto I-276 East (PA Turnpike) towards New Jersey. Continue East on I-276 (PA Turnpike) to Exit 339 Fort Washington. Go North on Route 309 (first right after going through tollgate). Follow 309 North approximately 3 miles to Norristown Road/Spring House exit. Go left onto Norristown Road for $\frac{3}{4}$ mile to McKean Road. Turn right on McKean Road and drive approximately $\frac{1}{2}$ mile to the J&J Spring House Campus entrance on your right side. *

PhilaSUG Executive Committee

Michael Davis, President
John Cohen, Membership Coordinator
Diane Foose, Treasurer
Robert Schechter, Newsletter Editor, and Web Master
Jonas V. Bilenas
Barry Cohen
Mark Keintz
Haibin Shu
Max Cherny
David Horvath
Steve Rhoades
Kajal Tahiliani

Host Sites Wanted

We continuously seek host sites for future PhilaSUG meetings. There is not a lot of work involved, and it is a great way to put your company on the local SAS map. We need your help with this. If your company would like to host a meeting, within reasonable geographic proximity to Philadelphia, PhilaSUG would be grateful if you would contact Michael Davis at michael.davis@alumni.duke.edu.

E-mail Announcements



PhilaSUG-L is a low volume, announcement-only e-mail notification service provided free of charge to all members who wish to subscribe. In order to sign up for this service, you need only send a blank e-mail message to PhilaSUG-L-subscribe@onelist.com. Note that you can subscribe as many times with as many different e-mail addresses as you wish to have the e-mail sent to; e.g., home and office.

PhilaSUG Web Site

Our site on the World Wide Web always contains the latest information concerning upcoming meetings, SAS training and seminars, links to SAS related hot topics, and local SAS job opportunities.



Visit us regularly at: <http://www.PhilaSUG.org>

Presenters Wanted



The PhilaSUG Executive Committee invites presentation abstracts. If you have thought presented before, PhilaSUG is an excellent venue for first-timers.our web site. Presentations can range from ten minute coders corner tips to 50 minutes talks. Both new and previously presented topics are welcome. If this is of interest to you, PhilaSUG invites you to submit electronic abstracts and bio for either a poster or oral presentation by using our online submission form. Check our home page for the Call for Papers.

A Thank You



The PhilaSUG Executive Committee wishes to thank Margie Merlino and Denis Michel at Janssen for arranging the hosting of this meeting. In addition, we wish to thank two members of our Executive Committee, Jonas Bilenas for speaker/program coordination and Michael Davis for site coordination.

The Present and Future of Define-XML

Lex Jansen, SAS Institute

Define-XML 2.1 is a metadata standard used to describe any tabular dataset structure. The primary use case for Define-XML is to describe CDISC Study Data Tabulation Model (SDTM), Standard for Exchange of Nonclinical Data (SEND), and Analysis Data Model (ADaM) datasets for the purpose of submissions to regulatory authorities. However, Define-XML also serves as a metadata exchange mechanism for other parties seeking to exchange CDISC-modeled dataset structures as well as proprietary (non-CDISC) dataset structures.

This presentation gives an overview of the current Define-XML 2.0 version, including Analysis Results Metadata. The presentation will also look ahead to the changes in the newest Define-XML 2.1 version.



Lex Jansen is a Principal Software Developer at SAS Institute, Health and Life sciences R&D. In this role, he develops software that supports data standards in the pharmaceutical industry.

Since 2008 Lex has been a member of the CDISC XML Technologies Team, where he has been active in the development of various CDISC standards: Define-XML 2.0/2.1, Dataset-XML and the Analysis Results Metadata extension for Define-XML 2.0.

Lex owns the website (www.lexjansen.com) which is well-known in the SAS community and contains more than 31,000 links to papers that were presented at major SAS User Group conferences.

§

For pharmaceutical company, oncology is one of the most important therapeutic area and for patients, oncology is one of the most unmet medical need. Oncology studies are very important and crucial to pharmaceutical industry. However, oncology studies are known as very complex and difficult. The presentation will show how programmers can lead oncology studies more effectively, easily and faster way.

First, the presentation will introduce the difference and specific aspects of oncology. It will show standards of each process - study types, response criteria, data collections, SDTM, ADaM, CT and analysis. The presentation will start with study types - Solid Tumor, Lymphoma and Leukemia, and its corresponding response criteria – RECIST, Cheson and IWCLL. The presentation will introduce oncology-specific data collections and corresponding CDISC standards implementation. The presentation will also introduce Time to Event ADaM and its corresponding oncology specific analysis – OS, PFS and Kaplan Meier Curve.

The presentation will show how standards of each step are linked in oncology studies from protocol, response criteria, data collection, SDTM, ADaM to analysis. The presentation will show how end to end standards driven process can help programmers and statistician to conduct effective, efficient, faster oncology studies.



Kevin Lee is Director of Data Science at Clindata Insight. Kevin always loves oncology studies and an active supporter on standards-driven process using CDISC and response criteria. Kevin has presented about 70 papers at the various conferences

including many related to oncology. Kevin earned an M.S. in Applied Statistics at Villanova University following a B.S. from University of Pennsylvania. Kevin is a life time learner who loves to learn and share.

§

**Developing Credit Risk Score Using SAS®
Programming**
Amos Odeleye, TD Bank

Credit Risk Score is an analytical method of modeling credit riskiness of individual borrowers (prospects and customers). While there are several generic, one-size-fit-all risk scores developed by vendors, there are numerous factors increasing driving the development of in-house Credit Risk Score. This presentation will introduce the audience on how to develop an in-house Credit Risk Score using SAS programming, Reject inference methodology and Logistic Regression.



Amos Odeleye graduated with B.S. in Statistics (University of Ibadan, NIGERIA), M.S. in Statistics (West Chester University, Pennsylvania) and M.A. in Economics (Temple University, Pennsylvania). Amos has more than 15 years of experience as a Statistician, Risk Manager, SAS programmer and as a consultant.

His current employment - VP, Credit Risk (TD Bank), and Adjunct Professor, Eastern University, St. Davids, Pennsylvania.

Amos' recent past employment - Assistant VP, Corporate Finance (TD Bank), Senior Statistician/Economist - AmeriHealth Caritas (a subsidiary of BlueCross BlueShield).

§

Proc Sort Revisited
Alex Chaplin, Bank of America

Proc sort can do more than sort your data. Revisit proc sort to see how you can select records and fields, rename and format fields, compress the output to save space, reuse space in your input dataset, remove and save off duplicate records and keys. Understand the difference between using the nodupkey and noduprecs options and how they affect your results at the aggregate and detail level.



Mr. Chaplin has been using SAS® since 2007. He is a SAS Certified Base and Advanced SAS Programmer. Mr. Chaplin has over 25 years of experience in Information Technology working a range of roles from programmer to technical project manager across a range of industries including travel, manufacturing, chemical, retail and financial in the UK, Spain and the US.

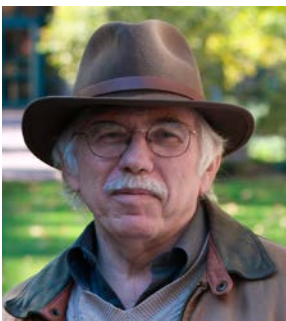
§

History Carried Forward, Mixing Time Series of Differing Frequencies

Mark Keintz, Wharton Research Data Services

Many programming situations require merging time series of varying frequency. For instance you might have three datasets (YEAR, QTR, and MONTH) of data, each with eponymous frequency and a common date variable. Producing a monthly file with the most recent quarterly and yearly data is a hierarchical last-observation-carried-forward (LOCF) task. Or you may have three irregular times series (ADMISSIONS, SERVICES, TESTRESULTS), in which you want to capture the latest data from each source at every date encountered (irregular LOCF).

This presentation demonstrates the use of conditional SET statements to update specific portions of the program data vector as a means of carrying forward low frequency data to multiple subsequent high frequency records, both for data sets with and without BY-groups. A similar approach works just as well for carrying forward data from irregular time series. There is also a demonstration how to use “sentinel variables” as a means of controlling the maximum length of time data is carried forward, i.e. removing data that has become “stale.”



Mark Keintz's current programming interests are in financial research applications and large data sets. He has previously served on NIH Small Business Innovation Research review panels and directed computing support for demography

and sociology research.

Mark has been using SAS® since it was documented in one book.

§

Using R to Query Dataset from SAS Server and Generate Forest Plots

Xiang (Jei) Li, Janssen Research and Development

The strength of SAS® is based on its powerful data manipulation capability, but it can be a hassle when using template language to generate customized plots, especially for beginners. On the other hand, R has exceptional flexibility in creating customized plots compared to SAS. However, the data management function of R tends to be time consuming and not as clean as SAS. In order to leverage the strength of each language, we will display how to integrate SAS dataset in R environment to generate forest plots.

Jei Li, after graduating with a MS degree in Biostatistics, he has been working at JNJ for two year as a Statistical Programming Analyst. Jei is now supporting Neuroscience TA and currently working on Esketamine TRD phase III compound.

§

Using SAS and Excel to Create Clear, Informative, and Realistic Mock Layouts in Word

Ken Maahs and Bob Zhong, Ph.D., Janssen Research and Development

Mock layouts are critical. They describe what programmers will create. They provide an initial picture of programming resource needs. They provide to the customer an early snapshot showing the planned content and are a key driver in helping to assess whether those plans are sufficient. Therefore, it is critical that mock layouts be clear, informative, realistic, and close to the look and feel of the final outputs. If they are not, outcome of the work is more likely to result in wasted time, rework, additional work, and additional resources being needed.

Mock layouts created in Word can often be challenging and time-consuming to create and change. If a suite of them needs an update all throughout, modifying each one can be a long and tedious process. As a result, we often rely on mock layouts that are not as clear and informative as they should be, and may even be symbolic in nature rather than an accurate image of what the customer can expect to see. To help with this problem, storing key mockup information in Excel and using SAS to process it into mock layouts will make creating and modifying those mockups a quick and easy process.

Ken Maahs, is a Principal Analyst (Senior Manager) in the Statistical Programming and Analysis functional area within Janssen Research and Development. Ken joined the company in 2007 as a Statistical Programmer Analyst. Initially, he supported the Immunology therapeutic area. After a couple years, Ken switched into Oncology, where he continue to provide support as a Therapeutic Area / Disease Area Expert. Ken earned a B.A. and M.S. in Mathematics from Eastern University (2004) and Drexel University (2006), respectively.

§

Automation of Clinical Trial Result Posting to ClinicalTrials.gov and EudraCT

Sherry Meeh, Janssen Research and Development

According to US FDAAA and EMA regulations, companies are now required to disclose their applicable clinical trial's aggregated results to US ClinicalTrials.gov and EMA EudraCT public websites: ClinicalTrials.gov (for US), and eudract.ema.europa.eu (for EMA). Currently in many pharmaceutical companies, the clinical trial aggregated results are prepared manually by a designated group based on the clinical study report. The manual process is time consuming and error-prone with iterative back-and-forth steps and a lot of reviews among different groups. Therefore, an innovative and automated process is needed to proactively streamline the result posting process, improve the accuracy, efficiency and consistency of the result posting, and reduce manual entry errors. In Johnson and Jonson, a biostatistics programming group, collaborating with biostatistics and clinical trial registry groups, developed an innovative automated process to proactively streamline the result posting process, improve the accuracy, efficiency and consistency of the result posting, and reduce manual entry errors by utilizing SAS.

Sherry Meeh has been a SAS user for over twenty years. She has worked in statistical programming group at pharmaceutical company, Contract Research Organizations and healthcare insurance company. Sherry is currently a Senior Manager of a statistical programming at Janssen. She has provided statistical programming support in the preparation of clinical trial reports, NDA/MAA submissions and result disclosures of clinical trial registry to clincialtrials.gov and EudraCT. She has extensive experience in variety of clinical trials. Sherry is also a member of PhUSE Data Transparency Working Group, and holds a Master of Science degree in Statistics.

§

Clinical Development Standards for FDA Bioresearch Monitoring (BIMO) Submissions

*Alice Jiang, Julie Maynard, and Denis Michel,
Janssen Research and Development*

The Food and Drug Administration (FDA) published the Bioresearch Monitoring Technical Conformance Guide in February 2018. The document provides specifications for clinical data submission by pharmaceutical companies used in planning of FDA Bioresearch Monitoring (BIMO) inspections. Three types of information are required: clinical study level information, subject level data line listings by clinical site, and a summary level clinical site dataset.

The clinical study level information are PDF files related to the clinical trial. The subject level data line listings by clinical site are typically generated as SAS programmed PDF files. The guide states that FDA will be able to generate the listings in the future from submitted clinical datasets compliant with CDISC SDTM and ADaM standards. The summary level clinical site dataset is provided as a SAS V5 transport file named clinsite.xpt with a data definition table named define.pdf. This presentation will describe the effort to standardize the generation of clinsite.xpt across different therapeutic areas of a pharmaceutical company. Topics include standardizing the SAS variable attributes via metadata, importing clinical site information from Excel to SAS, and developing standard SAS programs that process data in different formats.

Alice Jiang is a Senior Statistical Programmer Analyst at Janssen in Shanghai, China. Julie Maynard is a Standards Regulatory Intelligence Expert at Janssen in Pennsylvania, USA. Denis Michel is a Clinical Analysis Standards Lead at Janssen in New Jersey, USA. Alice, Julie, and Denis collaborate with a global Janssen Clinical Development Standards team developing internal Bioresearch Monitoring (BIMO) standards.

§

Future Meetings and SAS Events



PhUSE US Connect2019

Where: Baltimore Marriott Waterfront Hotel

When: February 24th - 27th 2019



Next PhilaSUG Meeting

To Be Announced

PhilaSUG Winter 2019 Meeting

Host: TBA

**When: TBA,
but targeting Mid-March
Through early April**



Where: Dallas, TX

When: April 28 - May 1

*Meeting Facility: Kay Bailey Hutchison
Convention Center*

Philadelphia Area SAS User Group (PhilaSUG) Membership Form

Membership only needs to be paid once per calendar year. Our membership year runs from Jan. 1st through Dec. 31st. Dues for the year are \$30. Membership entitles you to all PhilaSUG meetings free of charge. Your RSVP is all that's needed, otherwise please complete this form (please print) and return it to the registration desk of any PhilaSUG meeting (do NOT mail it). An attached business card helps legibility and is appreciated. This year we will be accepting cash, check (made out to "PhilaSUG"), or credit/debit card (this charge will show up on your credit/debit card statement as paid to "Michael Davis / PhilaSUG"). Please print.

Name: _____

Affiliation: _____

Day Time Phone Number: () _____ - _____ ← Only needed if paying by Credit/Debit Card

Privacy Statement - Local SAS User Groups are requested to share their membership/mailling list with SAS Institute on an annual basis. We respect your privacy and will never rent, sell or trade your personal information provided with any other group or individual and the information provided will only be used for PhilaSUG mailings. We will not share your email address with SAS unless you Opt In below.

Check this box (Opt In) if you agree to allow us to share *your email address with SAS*.

PhilaSUG-L is a low volume, announcement-only, e-mail notification service provided free of charge. By subscribing you'll be notified of the latest information about upcoming events, especially meeting announcements. By listing your e-mail address below you will be invited to join the electronic mailing list, you can cancel at any time.

E-mail:

@

provider i.e. yahoo gmail verizon comcast etc.

■ ← provider domain info i.e. com org edu net, etc.

(Please be sure to clearly distinguish a dash from an underscore)