

# 2014 Annual Conference of Pharmaceutical Industry SAS® Users Group

SAN DIEGO, CA June 1- 4, 2014



PROGRAM



**Susan Fehrer  
Coulson**  
Operations Chair  
BioClin, Inc.



**Jacques Lanoue**  
Academic Chair  
Novartis



## TABLE OF CONTENTS

---

Welcome to PharmaSUG 2014! .....	4
PharmaSUG 2014 Conference Committee.....	5
Operations Committee .....	5
PharmaSUG 2014 Section Chairs .....	6
Conference Schedule At-A-Glance .....	10
Conference Highlights.....	15
PharmaSUG 2014 Mobile App and Social Media .....	21
Opening Session & Mixer.....	23
Technical Keynote Session .....	23
SAS® Super Demos .....	24
PharmaSUG 2014 Single Day Events .....	28
SAS® at PharmaSUG 2014 .....	30
SAS® Papers and Seminars.....	31
<i>PharmaSUG CHINA</i> .....	32
PharmaSUG 2014 Training Seminars .....	33
<i>Exhibit &amp; Demo Room</i> .....	38
<i>Sapphire Ballroom</i> .....	39
<i>Indigo Ballroom</i> .....	40
<i>Cobalt Level</i> .....	41
<i>Aqua Level</i> .....	42
<i>Promenade</i> .....	43
Applications Development.....	44

Beyond the Basics ..... 53

Coders Corner ..... 59

Career Planning ..... 71

Data Visualization and Graphics ..... 75

Data Standards..... 79

Healthcare Analytics ..... 87

Hands-on-Training ..... 91

Industry Basics ..... 94

JMP..... 97

Management & Support ..... 98

Posters ..... 103

Statistics & Pharmacokinetics..... 109

Techniques & Tutorials: Foundations ..... 114



## WELCOME TO PHARMASUG 2014!

---

The Conference Committee of PharmaSUG 2014 welcomes you to the annual conference of the Pharmaceutical Industry SAS® Users Group! San Diego is a wonderful city in which to have the conference with the hotel right on the water with sunshine and warm temperatures. We are excited about some of the additions we have added to this year's event.

The conference kicks off on Sunday evening with an Opening Session featuring Mark Wolff of SAS Institute. He will speak about how his professional career changed after he presented at a PharmaSUG conference. His keynote address will address how solicited and non-solicited health-related issues information from the internet and social media can be harvested and verified to facilitate comparisons with clinical trials that are conducted in more controlled, scientifically rigorous environments.

Following the Opening Session, we will enjoy a mixer with food stations, drinks, and lots of socializing on the terrace. This mixer is a great opportunity to catch up with conference buddies, former colleagues, other SAS users, and relax before the paper presentations and Technical Keynote Address begin.

The Demo and Exhibit Rom will be open on Monday, Tuesday, and Wednesday morning. A new addition to the conference is a "Meet and Greet the Author" at the SAS Publishing area of the demo room. Authors of SAS books who are attending the conference will be available during the day to chat with you, the users. SAS Publishing has graciously offered to raffle off a book written by each author. Code doctors will again be available to answer your difficult programming questions. A small theater in the room will host the SAS Super Demos, Sponsor Presentations and our new section Career Planning. Please take the time to visit our exhibitors, including many SAS personnel who are available for questions and demos. Each day visit our exhibitor booths and get the chance to win a prize at our daily drawing.

We have several panel discussions planned thru the conference do not miss discussions, debate and prepare your questions. On Monday evening, we are fortunate to have the San Diego Padres ballpark, Petco Park, directly across the street, so Game Night has morphed into San Diego Padres game night. See the Padres play the Pittsburgh Pirates. We have reserved seating in the Kona Cove section, right at the third base line, for a very nominal fee. We will have ballpark fare: hot dogs, popcorn, peanuts, soft drinks, and water.

On Tuesday evening, a special SAS User Kickback Mixer awaits aboard the USS Midway Museum, a retired US Navy aircraft carrier permanently docked just one mile from the hotel. Bus service to and from the mixer will be provided. We will have docents to lead ship tours, a roving photographer, flight simulators, food, and drinks. Please be sure to wear comfy shoes! If you choose to walk back to the hotel, it is a lovely walk along the waterfront through a little village of shoppes and restaurants.

Before and after the conference, we will again be offering a variety of seminars taught by renowned experts. Seminars will be offered on Saturday and Sunday, morning and afternoon, Wednesday afternoon, and all day Thursday. We will also be offering a discounted rate for the SAS certification exams on Wednesday afternoon.

**We wish you a great conference that is educational and fun. Welcome to San Diego!!!!**

**Susan Fehrer Coulson**  
Operational Chair

**Jacques Lanoue**  
Academic Chair

## PHARMASUG 2014 CONFERENCE COMMITTEE

### Conference Chairs

#### Operations Chair

Susan Fehrer Coulson  
BioClin, Inc.

#### Academic Chair

Jacques Lanoue  
Novartis

### Operations Committee

Registrar	Elizabeth Dennis	EMB Statistical Solutions
Assistant Registrar	Eric Larson	Ephicity Consulting Group
Registration Desk Coordinator	Elaine Dempsey	
Digital Media Coordinator	Syamala Schoemperlen	Alcon
Meals Coordinators	Ellen Brookstein	Philadelphia University SAS Certification Program
	Syamala Schoemperlen	Alcon
Seminar Coordinators	Sandra Minjoe	Accenture
	Richann Watson	Experis
A/V Coordinator	Margaret Hung	MLW Consulting
Volunteer Coordinators	Craig Brelage	Eli Lilly and Co.
	Lynn Mullins	
Scholarship Coordinators	Richard Allen	Peak Statistical Services
	Ellen Brookstein	Philadelphia University SAS Certification Program
Sponsor & Exhibitor Coordinators	Cindy Song	Sanofi
	Paul Slagle	inVentiv Health Clinical
Speaker Sharing Coordinators	Richard Allen	Peak Statistical Services
	Ellen Brookstein	Philadelphia University SAS Certification Program
Padres Game Night Coordinator	Kim Truett	KCT Data, Inc.
Speaker Training Coordinator	Mal Foley	University of North Carolina
First Timers' Coordinators	Ellen Brookstein	Philadelphia University SAS Certification Program
	MaryAnne DePesquo	BCBS of Arizona
Media & Materials Coordinators	Josh Horstman	Nested Loop Consulting
	Natalie Reynolds	Eli Lilly and Co.
Philanthropy Coordinator	Shannon Larson	
Photographers	James Wu	Mitsubishi Tanabe Pharma
	Gary Moore	Moore Computing Services, Inc.
FDA & Industry Recruiter	Dave D'Attilio	TalentMine
Copyright Coordinator	Lin Yan	Celgene Corp.
Webmaster	Josh Horstman	Nested Loop Consulting

## PHARMASUG 2014 SECTION CHAIRS

---

Applications Development	Peter Eberhardt	Fernwood Consulting Group
	Gary Moore	Moore Computing Services, Inc.
Beyond the Basics	Kirk Lafler	Software Intelligence Corp.
	Xiaohui Wang	Novartis
Career Planning	Josh Horstman	Nested Loop Consulting
	Marlene Miller	Ockham
Code Doctors	John King	Ouachita Clinical Data Services
Coders Corner	Mal Foley	University of North Carolina
	David Franklin	TheProgrammersCabin.com
Data Standards	Rob Howard	Veridical Solutions
	Mario Widel	Eli Lilly and Co.
Data Visualization and Graphics	John Labore	Eli Lilly and Co.
	Nina Worden	PPD
Hands-On Training	Nancy Brucken	inVentiv Health Clinical
	Alissa Ruelle	inVentiv Health Clinical
Healthcare Analytics	Richard Allen	Peak Statistical Services
	MaryAnne DePesquo	BCBS of Arizona
Industry Basics	Meera Kumar	M & G Insights
	Carol Matthews	United Biosource Corporation
Management and Support	Jim Baker	Cytel
	Kent Letourneau	PRA International
Posters	Amie Bissonett	inVentiv Health Clinical
	Subrahmanyam (Gopal) Rajagopal	TAKE Solutions, Inc.
Statistics and Pharmacokinetics	Vincent Amoruccio	Alexion Pharmaceuticals
	Michael Wilson	IU School of Medicine
Techniques & Tutorials: Foundations	Cecilia Mauldin	Assent
	Steven Michaelson	SimulStat

Let Ockham **Source** help you  
**STEER YOUR**  
**OUR**  
**CAREER**

Helping you navigate your professional path

Ockham Source is proud to announce our **Diamond Sponsorship** of this year's PharmaSUG conference. We'd like to invite you to visit us at Booth #1 to discover how we can help you navigate through your professional career. Our niche market expertise, industry experience, career coaching, and much more, can help you experience true career success.

 **We're SOCIAL**

Follow us on Twitter for regular conference updates during PharmaSUG.

@Ockham\_Source

*Driven by outcome. Defined by expertise.*

**OCKHAM**  
SOURCE

[WWW.OCKHAM.COM](http://WWW.OCKHAM.COM)



VISIT US AT BOOTH #1



## Steer Your Career with Ockham Source

Ockham Source, a full-spectrum global resourcing solution, is here to help navigate you through your professional career. Let our experts find the best fit for you and the professionals with whom we partner. Together we can make your career path a success.

### Resources for Talent

- Career Search
- Career Development
- Global Opportunities

### Resources for Clients

- Specialized Expertise
- Functional Service Provider (FSP)
- Contract to Permanent Placements

*Driven by outcome. Defined by expertise.*

PharmaSUG Diamond Sponsor  
[WWW.OCKHAM.COM](http://WWW.OCKHAM.COM)

Pharma  
**SUG**  
SAN DIEGO 2014



# Thank You!

## To Our Sponsors

DIAMOND



PLATINUM



# GILEAD

GOLD



SILVER



BRONZE



## CONFERENCE SCHEDULE AT-A-GLANCE

Time	Event	Location
<b>Saturday, May 31</b>		
7:30am - 8:00am	Seminar Registration	<i>Indigo West Foyer</i>
8:00am - 5:00pm	<a href="#">Pre-conference Training Seminars</a> (separate fee)	<i>Indigo C-D-H</i>
<b>Sunday, June 1</b>		
7:30am - 8:00am	Seminar Registration	<i>Indigo West Foyer</i>
8:00am - 5:00pm	<a href="#">Pre-conference Training Seminars</a> (separate fee)	<i>Indigo C-D-H</i>
12:30pm - 1:00pm	Seminar Registration	<i>Indigo West Foyer</i>
1:00pm - 6:30pm	Conference Registration Desk Open	<i>Sapphire West Foyer (4th Floor)</i>
1:30pm - 5:00pm	Presentation Rehearsal	<i>Indigo C</i>
2:00pm - 4:00pm	Public Speaking & Presentation Training	<i>Sapphire M</i>
4:00pm - 5:00pm	First Timers' Session	<i>Sapphire M</i>
5:00pm - 6:00pm	Presenter, Section Chair, and Volunteer Meeting	<i>Sapphire M</i>
6:30pm - 7:30pm	<a href="#">Opening Session</a>	<i>Indigo ABEF</i>
7:30pm - 11:00pm	<a href="#">Opening Mixer</a>	<i>Indigo ABEF</i>
<b>Monday, June 2</b>		
7:00am - 8:00am	Continental Breakfast	<i>Sapphire West Foyer</i>
7:00am - 8:00am	Student Scholar & Jr. Professional Award Recipient Breakfast	<i>Aqua 300</i>
7:30am - 8:00am	Conference Registration Desk Open	<i>Sapphire West Foyer</i>
8:00am - 8:45am	<a href="#">Technical Keynote Session</a>	<i>Indigo ABEF</i>
9:00am - 10:00am	Conference Registration Desk Open	<i>Sapphire West Foyer</i>
9:00am - 11:30am	Code Doctors	<i>Demo Room and Exhibit Area</i>
9:00am - 11:45am	<a href="#">Demo Room and Exhibit Area</a> Open	<i>Sapphire Ballroom</i>
9:00am - 11:45am	SAS Super Demos	<i>Demo Theater (Exhibit Area)</i>
9:00am - 12:05pm	<a href="#">Paper Presentations and Hands-On Training</a>	<i>Sapphire Level</i>
10:00am - 10:15am	Morning Coffee Break	<i>Demo Room and Exhibit Area</i>
12:00pm - 1:15pm	Networking Lunch (included)	<i>Indigo ABEF</i>
12:45pm - 6:00pm	<a href="#">Demo Room and Exhibit Area</a> Open	<i>Sapphire Ballroom</i>
1:15pm - 5:30pm	<a href="#">Paper Presentations and Hands-On Training</a>	<i>Sapphire Level</i>
1:30pm - 4:30pm	SAS Press Author Meet & Greet	<i>Demo Room and Exhibit Area</i>
1:30pm - 5:00pm	Code Doctors	<i>Demo Room and Exhibit Area</i>
2:00pm - 4:00pm	Conference Registration Desk Open	<i>Sapphire West Foyer</i>
2:30pm - 3:30pm	Meet the Poster Authors	<i>Sapphire West Foyer</i>
3:15pm - 3:30pm	Afternoon Refreshment Break	<i>Demo Room &amp; Exhibit Area</i>
6:10pm - 10:00pm	<a href="#">Padres Game Night</a> (separate fee)	<i>Petco Park (across street)</i>
7:00pm - 11:00pm	<a href="#">Scriptathon</a>	<i>Sapphire P</i>

Time	Event	Location
<b>Tuesday, June 3</b>		
7:00am - 8:00am	Continental Breakfast	<i>Indigo ABEF</i>
8:00am - 10:00am	Conference Registration Desk Open	<i>Sapphire West Foyer</i>
8:00am - 12:05pm	<a href="#">Paper Presentations and Hands-On Training</a>	<i>Sapphire Level</i>
9:00am - 11:30am	Code Doctors	<i>Demo Room and Exhibit Area</i>
9:00am - 11:45am	<a href="#">Demo Room and Exhibit Area</a> Open	<i>Sapphire Ballroom</i>
9:00am - 11:45am	SAS Super Demos	<i>Demo Theater (Exhibit Area)</i>
10:00am - 10:15am	Morning Coffee Break	<i>Demo Room and Exhibit Area</i>
12:00pm - 1:15pm	Networking Lunch (included)	<i>Indigo ABEF</i>
12:45pm - 6:00pm	<a href="#">Demo Room and Exhibit Area</a> Open	<i>Sapphire Ballroom</i>
1:15pm - 5:30pm	<a href="#">Paper Presentations and Hands-On Training</a>	<i>Sapphire Level</i>
1:30pm - 5:00pm	Code Doctors	<i>Demo Room and Exhibit Area</i>
2:00pm - 4:30pm	SAS Press Author Meet & Greet	<i>Demo Room and Exhibit Area</i>
2:00pm - 4:00pm	Conference Registration Desk Open	<i>Sapphire West Foyer</i>
2:30pm - 3:30pm	Meet the Poster Authors	<i>Sapphire West Foyer</i>
3:15pm - 3:30pm	Afternoon Refreshment Break	<i>Demo Room &amp; Exhibit Area</i>
6:00pm - 6:30pm	Bus pickup for <a href="#">SAS User Appreciation Mixer</a>	<i>1st Floor / Promenade East</i>
6:30pm - 9:30pm	<a href="#">SAS User Appreciation Mixer aboard U.S.S. Midway</a> (included)	<i>U.S.S. Midway Museum</i>
<b>Wednesday, June 4</b>		
7:00am - 8:00am	Continental Breakfast	<i>Indigo ABEF</i>
8:00am - 10:00am	Conference Registration Desk Open	<i>Sapphire West Foyer</i>
8:00am - 11:05am	<a href="#">Paper Presentations and Hands-On Training</a>	<i>Sapphire Level</i>
8:00am - 11:00am	<a href="#">Demo Room and Exhibit Area</a> Open	<i>Sapphire Ballroom</i>
8:00am - 10:45am	SAS Super Demos	<i>Demo Theater (Exhibit Area)</i>
9:00am - 11:00am	Code Doctors	<i>Demo Room and Exhibit Area</i>
10:00am - 10:15am	Morning Coffee Break	<i>Demo Room and Exhibit Area</i>
11:30am - 12:30pm	<a href="#">Closing Session</a>	<i>Sapphire DHL</i>
12:30pm - 1:00pm	Seminar Registration	<i>Cobalt Lobby</i>
1:00pm - 5:00pm	<a href="#">Post-conference Training Seminars</a> (separate fee)	<i>Cobalt 500, 501, 502</i>
1:00pm - 5:00pm	<a href="#">SAS Certification Exams</a> (separate fee)	<i>Sapphire P</i>
<b>Thursday, June 5</b>		
8:00am - 5:00pm	<a href="#">Post-conference Training Seminars</a> (separate fee)	<i>Cobalt 500, 501, 502</i>



*Thank You To Our Exhibitors*





*Thank You To Our Exhibitors*





Gilead was founded in 1987 in Foster City, California. In 25 years, Gilead has become a leading biopharmaceutical company with a rapidly expanding product portfolio, a growing pipeline of investigational drugs and approximately 6,400 employees in offices across four continents. Millions of people around the world are living healthier, more fulfilling lives because of innovative therapies developed by Gilead. To our research and development effort is the largest it has ever been, with more than 130 clinical studies evaluating compounds with the potential to become the next generation of innovative therapies for HIV/AIDS, liver diseases, serious respiratory and cardiovascular conditions, cancer and inflammation. In line with our continued growth, Gilead Sciences is seeking qualified candidates for the following positions. The roles are based in our corporate headquarters in Foster City, CA (San Francisco Bay Area) including our Seattle, WA site.

**Sr. Manager, Statistical Programming - Req. #13660**  
Seeking a hands-on manager to mentor and develop the skills of an experienced team.

**Manager, Statistical Programming - Req. #13693**  
Manages and motivates the team of statistical programmer supporting Gilead's Oncology therapeutic area.

**Senior Statistical Programmer - Req. #16056**  
Works as a lead programmer for multiple Phase studies and as a secondary programming contact for one or more Phase 1-3 studies.

**Statistical Programming II - Req. #15057**  
Works as secondary programming contact for 2-3 Phase 1-3 studies



# GILEAD

Advancing Therapeutics.  
Improving Lives.

Come visit us on the Exhibit  
Floor - Booth 9

<http://biometrics-careers.gilead.com>

## CONFERENCE HIGHLIGHTS

---

### **Presentation Rehearsal**

Sunday, Jun 1 - 1:30 PM to 5:00 PM

Indigo C

A presentation rehearsal room is available for presenters to practice their presentations. Prior sign-up outside the door is required.

### **Public Speaking & Presentation Training**

Sunday, Jun 1 - 2:00 PM to 4:00 PM

Sapphire M

Presenters, PharmaSUG is your chance to shine! Maximize the impact of your time in the spotlight by attending this free training session on giving effective presentations. You'll have the opportunity to hone your skills with the expert guidance of seasoned speakers

### **First Timers Session**

Sunday, Jun 1 - 4:00 PM to 5:00 PM

Sapphire M

Is it your first time at PharmaSUG? Wonder what to see and how to fit it all in? Would you like to win some door prizes? Experienced conference organizers and our SAS liaison will provide you with tips on selecting presentations, advice on managing your time, and an overview of what and where you can find information in the Exhibit & Demo Room. You'll also have the chance to win fabulous door prizes!

### **Presenter, Section Chair, & Volunteer Meeting**

Sunday, Jun 1 - 5:00 PM to 6:00 PM

Sapphire M

All paper presenters, section chairs, and volunteers are strongly urged to attend the Presenter, Section Chair, and Volunteer meeting on Sunday afternoon. Presenters and volunteers will meet with their section chairs and receive important instructions to ensure the smooth operation of our conference.

### **Opening Session & Keynote**

Sunday, Jun 1 - 6:30 PM to 7:30 PM

Indigo AB EF

The 2014 conference officially gets underway on Sunday evening with our opening session. You'll have a chance to meet the conference co-chairs, learn about this year's charity fundraiser, and hear our keynote address from Mark Wolff of SAS Institute. Mark's presentation is entitled "Post-Marketing Safety Monitoring of Self-Reported Symptom-Treatment-Outcome Measurement in Web-based Media Sources". This is sure to be of interest to anyone who has ever diagnosed themselves using the Internet! Afterwards, we'll enjoy a mixer with food stations, drinks, and plenty of socializing. This will be a great opportunity to meet other attendees, network with fellow SAS users, and just relax before we dig into the technical content on Monday.

### **Opening Mixer**

Sunday, Jun 1 - 7:30 PM to 11:00 PM

Indigo Foyer and Patio

We'll enjoy a mixer with food stations, drinks, and plenty of socializing. This will be a great opportunity to meet other attendees, network with fellow SAS users, and just relax before we dig into the technical content on Monday.

### Technical Keynote Session

Monday, Jun 1 - 8:00 AM to 8:45 AM

Indigo ABEF

PharmaSUG is pleased to welcome Jack Shostak, Associate Director of Statistics at Duke Clinical Research Institute. Jack will present our Monday morning technical keynote session on The Future of SAS Programming in the Pharmaceutical Industry. Needless to say, if you are a SAS programmer and would like to continue that work, you won't want to miss out on Jack's insights about the future of the industry.

### Padres Game Night

Monday, Jun 2, 6:10 PM – 10:00 PM

Petco Park, San Diego Padres, Arena/Stadium

On Monday evening, we are fortunate to have the San Diego Padres ballpark, Petco Park, directly across the street, so Game Night has morphed into San Diego Padres game night. See the Padres play the Pittsburgh Pirates. We have reserved seating in the Kona Cove section, right at the third base line, for a very nominal fee. We will have ballpark fare: hot dogs, popcorn, peanuts, soft drinks, and water.

### PharmaSUG/PhUSE Scriptathon

Monday, Jun 2, 7:00 PM – 11:00 PM

Sapphire Ballroom P

Building on the momentum of the March 2014 PhUSE Computer Science Symposium Scriptathon, PharmaSUG and PhUSE are collaborating on a second Scriptathon to enhance the PhUSE Code repository. Scripts (SAS code) will be developed to produce various types of standard data displays, or targets. Come join us from 7:00-11:00 PM on Monday, June 2, in Sapphire P (Hands-On Training room), and help in the effort to build a usable code library! For this event, targets will be culled from the [CTSpedia collection](#), and scripts developed using the CDISC ADaM pilot project datasets. All materials will be provided. Please note that you must be registered for the conference in order to participate.

More information on the PhUSE CSS Standard Scripts Working Group can be found on the [PhUSE Wiki](#).

### Code Doctors

Monday, May 13 9:00 AM to 11:30 AM & 1:30 PM to 5:00 PM

Demo Theater - Exhibit & Demo Room

Tuesday, May 14 9:00 AM to 11:30 AM & 1:30 PM to 5:00 PM

Demo Theater - Exhibit & Demo Room

Wednesday, May 15 9:00 AM to 11:00 AM

Demo Theater - Exhibit & Demo Room

Have a free consultation with our Code Doctors. Bring a printout and/or e-file containing your code, source data and error messages for a definitive diagnosis. Please check the PharmaSUG Mobile App or the posted schedule in the Exhibit & Demo Room for when the "Doctors are In!!"

### Demo Room and Exhibit Area

Monday, Jun 2 9:00 AM to 11:45 AM & 12:45 PM to 6:00 PM

Sapphire BCFGJKNO

Tuesday, Jun 3 9:00 AM to 11:45 AM & 12:45 PM to 5:00 PM

Sapphire BCFGJKNO

Wednesday, Jun 4 8:00 AM to 11:00 AM

Sapphire BCFGJKNO

The Exhibit & Demo Room will be open Monday through Wednesday and will feature an extensive display of SAS products and services. SAS Staff will be on hand to discuss and demonstrate the latest software and services available.

In addition, you'll find a variety of booths hosted by exhibitors wanting to connect with you and other attendees. This year's exhibitors include contract research organizations (CROs), staffing and recruiting companies, university degree programs, software tool vendors, and more. (Only registered exhibitors are permitted to display their lit-

erature and products.) There is also a PharmaSUG table with information about upcoming conferences. A job position posting folder will be available at the registration desk.

### **Hands-On Training**

Sapphire P

We are once again pleased to offer our own Hands-on Training Section, this year in a classroom-style setting. Presenters were invited to participate, and SAS Institute provides the computers for our use. Abstracts for the Hands-On Training are listed in the Abstracts section (Separate document).

### **Posters**

Monday, Jun 2, 2:30 PM – 3:30 PM

Poster Sapphire West Foyer

Tuesday, Jun 3, 2:30 PM – 3:30 PM

Poster – Sapphire West Foyer

Posters are another way to present technical information for viewing. These are displayed Monday and Tuesday for your viewing pleasure and learning. Abstracts for the Posters are also listed in this program book. Each poster also includes a paper that will be posted as part of the PharmaSUG online proceedings at [www.pharmasug.org](http://www.pharmasug.org). Stop by during Meet the Poster Authors from 2:30 PM to 3:30 PM on Monday and Tuesday to meet one-on-one with the poster authors.

Don't forget to stop and view the award winning posters from various regional SAS user group conferences and SAS Global Forum.

### **Closing Session**

Wednesday, Jun 4, 11:30 AM - 12:30 PM

Sapphire HDL

Leaving so soon? Think again! Wrap up your PharmaSUG conference at our Closing Session. We'll announce the Best Paper winners, introduce next year's conference co-chairs, and of course give out **TONS OF DOOR PRIZES!**

### **Position Referral Manual**

A Position Referral Manual with job postings and resumés will be on display during the conference at the PharmaSUG table next to the Exhibit & Demo Room. There is no fee for placing job postings or resumés in the binder, but the material must fit into an 8.5" x 11" clear plastic sheet protector. You are strongly encouraged to place multiple copies of your material in the sheet protector so interested attendees may take one. We reserve the right to remove inappropriate material. *Direct recruiting at the conference is prohibited.*

### **Meals**

Your conference registration fee includes the following meals: Sunday night mixer plus lunch on Monday and Tuesday. A light continental breakfast on Monday, Tuesday, and Wednesday mornings is also included in your conference registration fee.

### **Twitter Give away**

If you follow us @pharmasug on Twitter or on the Twitter feed in the PharmaSUG Mobile App, you will receive instructions on how PharmaSUG and the Exhibitors will be giving away additional prizes.

### **Daily Drawings and Tweet-Aways**

Win great door prizes like tablet devices, gift cards, and more in each of our daily drawings. Simply visit a certain number of vendor booths each day to qualify for entry into the drawing! Details will be included in your conference registration materials. We will also be giving away additional prizes by Tweet-away. If you follow us @pharmasug on Twitter or on the Twitter feed in the PharmaSUG Mobile App, you'll receive instructions on how to win!

### **Charity Drive**

In 2013 we raised over \$3700 for STEM education at a downtown Chicago high school. While PharmaSUG continues to support STEM education, literacy is a much larger concern in San Diego (and much of California), and without the ability to read, students will not have opportunities to explore STEM or other kinds of educational paths. So, this year our goal is to raise \$4000 for a San Diego middle school classroom (7th and 8th grade) struggling with illiteracy. Perkins K-8 is located less than a mile southeast of our hotel in the Logan Barrio. This particular school has some notable statistics:

- Primary ethnicity is Hispanic/Latino (95%), followed by African American (3%)
- 57% are English learners
- 99% are eligible for free/reduced lunch
- 71% of parents have a high school education or less

Classroom URL: [http://www.adoptaclassroom.org/classroomdonation/results\\_teacher.aspx?classroomid=175268](http://www.adoptaclassroom.org/classroomdonation/results_teacher.aspx?classroomid=175268)

**Attendees will be able to make a donation using cash, check, or by credit card through the Adopt-a-Classroom website.**

### **Badges**

Your name badge will be your admission to the entire conference. It is required to attend the presentations and meal functions. You must wear your name badge in a visible location during all conference activities.

### **PharmaSUG Web Site**

[www.PharmaSUG.com](http://www.PharmaSUG.com) PharmaSUG 2014 is the premier event of the year for SAS users in the pharmaceutical, healthcare, and related industries!

Join us for two and a half days of paper presentations, networking events, demonstrations and exhibits, hands-on training, and much more! Enhance your experience with pre- and post-conference training seminars.

After the conference, visit the PharmaSUG web site at [www.pharmasug.org](http://www.pharmasug.org) to:

1. access presented papers
2. read a wrap-up and review of the conference
3. complete the on-line conference survey
4. find out the best paper award winners by sectionsave the date and location for next year's PharmaSUG conference

# Strategic Resourcing

- »» **Enhanced Productivity**
- »» **Managed Costs**
- »» **Increased Value**

Offering the most experienced strategic resourcing team in the industry, inVentiv Health Clinical offers innovative, customizable, cost-effective solutions from contingent staffing to functional models to strategic partnerships to help improve quality, increase efficiency, and manage costs.

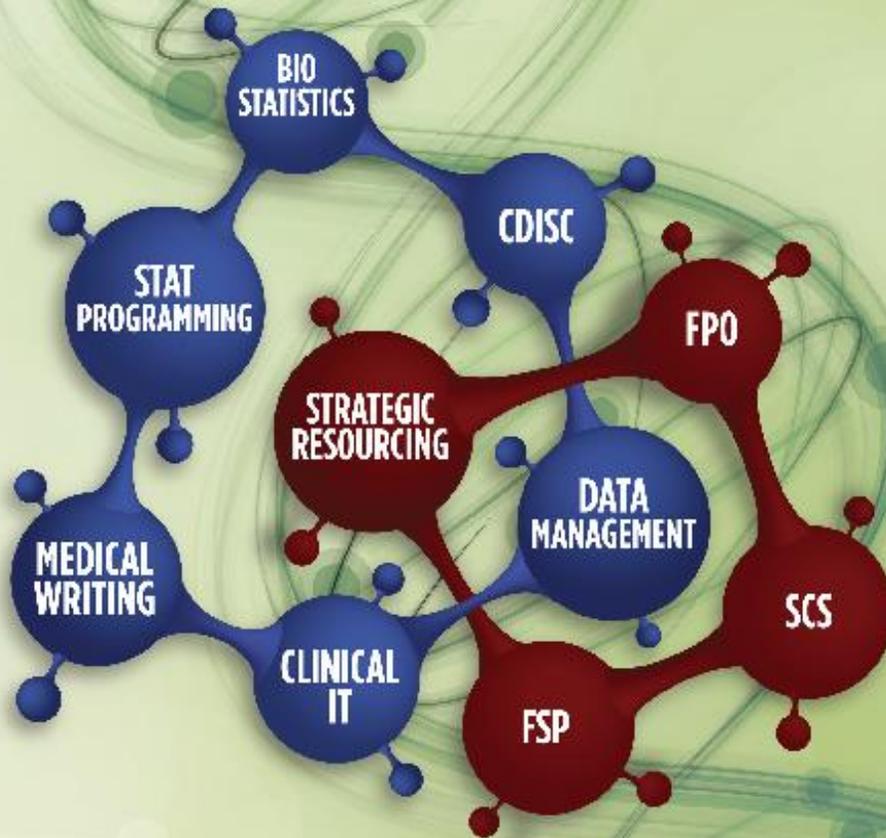
**To learn how to increase the value of your program, visit us at [www.inVentivHealthclinical.com](http://www.inVentivHealthclinical.com).**

»» **Please visit us at Booth #11**



[www.inVentivHealthclinical.com](http://www.inVentivHealthclinical.com)

*Your Partner for Excellence...*



See us at BOOTH#31 @PharmaSUG 2014



**GCE SOLUTIONS**

*Derive Value from Excellence...*

**Experts at delivering Clinical Data Services**

For more Info: Visit us at [www.gcesolutions.com/cds](http://www.gcesolutions.com/cds) or email us at [cds@gcesolutions.com](mailto:cds@gcesolutions.com)

## PHARMASUG 2014 MOBILE APP AND SOCIAL MEDIA

---



### PharmaSUG 2014 Digital Media Guide

In an effort to reduce paper waste, this year's conference features a virtual program guide. All of the information previously found in a printed program can be accessed in one of the following ways:

#### PHARMASUG MOBILE APP

The PharmaSUG Mobile App is your digital guide to the PharmaSUG 2014 conference. It includes detailed schedule information, presentation abstracts, maps, information about sponsors and exhibitors, last-minute updates, and much more.

- iPhone/ iPad: search "pharmasug 2014" on the Apple App Store.
- Android devices: search "pharmasug 2014" on the Google Play Store.
- Blackberry, Windows Phone, and others: Access the mobile app by pointing your mobile browser to the following address: <https://pharmasug2014.gatherdigital.com>
- You can even use the address above to view the content in a traditional browser on your laptop/desktop.

#### PDF CONFERENCE PROGRAM

A PDF version of the conference program and paper abstracts are available for download from the PharmaSUG website (<http://pharmasug.org/us/2014/mobile.html>).

Prefer a hard copy? Printed copies of the conference program and paper abstracts are available at the UPS Store on the 3<sup>rd</sup> floor of the hotel for a nominal fee.

### Social Media at PharmaSUG 2014



#### TWITTER GIVE AWAY

Follow us @pharmasug on Twitter or on the Twitter feed in the PharmaSUG Mobile App to receive instructions on how PharmaSUG and the Exhibitors will be giving away additional prizes. We encourage you to tweet about your conference experience through the event!



#### LINKEDIN

Stay connected to the people you have met at the conference! Join the PharmaSUG LinkedIn group: <http://www.linkedin.com/groups?gid=2006675>



THE BEST EXECUTIVE DECISION

POSSIBLE™

Teradata is the world's leading analytics solutions company. Focused on improved data warehousing, big data analytics, and business applications. We empower organizations to make the best decisions possible for competitive advantage.

**TERADATA.**

[teradata.com](http://teradata.com)

## OPENING SESSION & MIXER

---

**Featuring SAS Institute's Mark Wolff**  
**Sunday, June 1, 6:30pm - 11:00pm**



The 2014 conference officially gets underway on Sunday evening with our opening session. You'll have a chance to meet the conference co-chairs, learn about this year's charity fundraiser, and hear our keynote address from Mark Wolff of SAS Institute. Mark's presentation is entitled "Post-Marketing Safety Monitoring of Self-Reported Symptom-Treatment-Outcome Measurement in Web-based Media Sources". This is sure to be of interest to anyone who has ever diagnosed themselves using the Internet! Afterwards, we'll enjoy a mixer with food stations, drinks, and plenty of socializing. This will be a great opportunity to meet other attendees, network with fellow SAS users, and just relax before we dig into the technical content on Monday.

**Mark Wolff, Ph.D.**  
**Principal Industry Consultant**  
**Health & Life Sciences Global Practice**  
**SAS Institute**

## TECHNICAL KEYNOTE SESSION

---

**"The Future of SAS Programming in the Pharmaceutical Industry"**  
**Featuring Jack Shostak of Duke Clinical Research Institute**  
**Monday, June 2, 8:00am - 9:00am**



PharmaSUG is pleased to welcome Jack Shostak, Associate Director of Statistics at Duke Clinical Research Institute. Jack will present our Monday morning technical keynote session on The Future of SAS Programming in the Pharmaceutical Industry. Needless to say, if you are a SAS programmer and would like to continue that work, you won't want to miss out on Jack's insights about the future of the industry.

**Jack Shostak**  
**Associate Director of Statistics**  
**Duke Clinical Research Institute**

# Biostatistics & Programming Outsourcing

Boosting Merck Serono's productivity with a Cytel "extension arm"

**Challenge:** Assemble a biostatistics and clinical programming team qualified to synchronize with Merck Serono counterparts

**Benefit:** Increased clinical data programming, analyses and reporting at lower costs

## Sponsor's Challenge

Biostatistics and statistical programming personnel are hard-pressed to keep up in today's "do more with less" biopharma industry environment.

Why? Typically, teams must support multiple development programs. They must also contend with growing product pipelines stemming from improved discovery efforts, increased early stage clinical successes, and programs inherited from mergers & acquisitions.

Even the most talented and efficient teams strain to keep pace with ever more clinical trials underway and on the horizon. Many feel they're "victims of our own success."

Is outsourcing a viable, sustainable option?  
Can outside biostatisticians and programmers sufficiently boost the sponsor's own stats and SAS programming output without compromising quality?

Merck Serono recently put that very question to Cytel. Key for the European-headquartered biopharmaceutical was sufficient demonstration of 1. the remote personnel's proficiency levels, and 2. the capacity to maintain accuracy even despite anticipated scaling.



**Securing the project with Merck Serono required three essential components:**

1. Finding a team qualified to work with Merck Serono personnel
2. Meeting Merck Serono standards, in adherence with FDA & EMA regulations
3. A lower cost solution than internal hiring



Cytel Clinical Services Case Study 5

# SAS® SUPER DEMOS

SAS Super Demos

## 2014 SAS Super Demos

SAS Super Demos are live 30-minute software demonstrations provided by experts from SAS Institute. Super Demos will run through each morning of the conference in our Demo Theater and cover a variety of topics. The Demo Theater is located at the back of the Demo Room and Exhibit Area in Sapphire N. Seating is limited, so come early!

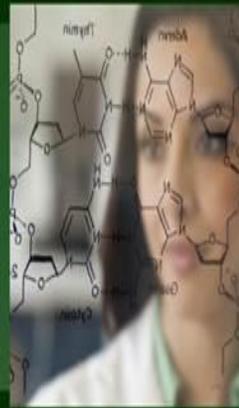
Please note that this schedule is subject to change. Last update: 09-May-2014.

Time	Presentation	Presenter
<b>Monday, June 2</b>		
9:00 - 9:30am	SAS Clinical Data Integration: How to Develop a Custom Generated Transformation	Melissa Martinez
9:30 - 10:00am	SAS Solutions OnDemand Extensions for SAS Drug Development (SDD) - An Example of SDD API Usage	Melissa Martinez
10:15 - 10:45am	Define.xml Round-Trip in SAS Clinical Data Integration	Julie Maddox
10:45 - 11:15am	The Integration of SAS Drug Development and SAS Clinical Data Integration	Julie Maddox
11:15 - 11:45am	Visualizing Clinical Data with SAS Visual Analytics	Matt Becker
<b>Tuesday, June 3</b>		
9:00 - 9:30am	What's New in SAS/STAT 13.1	Maura Stokes
9:30 - 10:00am	SAS Clinical Data Integration: Turning Your Data Around	Janet Stuelpner
10:15 - 10:45am	Clinical Graphs using SG Procedures	Sanjay Matange
10:45 - 11:15am	What's new in ODS Graphics for SAS 9.4	Sanjay Matange
11:15 - 11:45am	The SAS Drug Development Audit History and Enhancements Using Extended Attributes	Melissa Martinez
<b>Wednesday, June 4</b>		
8:00 - 8:30am	Creating Dataset-XML files with the SAS Clinical Standards Toolkit	Lex Jansen
8:30 - 9:00am	Creating Define-XML Version 2 with the SAS Clinical Standards Toolkit 1.6	Lex Jansen
9:00 - 9:30am	Creating Portable Jobs in SAS Drug Development	Sharon Trevoy
9:30 - 10:00am	SAS Drug Development API	Sharon Trevoy
10:15 - 10:45am	Modifying the PROC LIFETEST Survival Plot	Warren Kuhfeld



# ALPHA CONSULTING CORP.

The Pinnacle of Staffing Solutions



# CRC Pharma LLC

*A clinical reporting company  
for the pharmaceutical industry*



## About us

CRC Pharma is a Contract Research Organization focused on the statistical reporting of clinical drug trials. We also provide related services.

Ramesh Maddena, President, has more than fourteen years experience as a clinical programmer and technical leader with various medium and large pharmaceutical companies. He holds a BS in Mathematics and a MS in Computer Applications from Nagarjuna University, India.

William Stewart, PhD, Vice-president, has over thirty years experience in the pharmaceutical industry in all phases of clinical development as a hands-on statistician, statistical group head, statistical quality group head, and a statistical reporting group head. He holds a MS in Statistics from Florida State University and a PhD in Psychometrics and Research Design from City University of New York.

Both Mr. Maddena and Dr. Stewart have extensive experience in selecting, managing and evaluating large and small CROs. They are intimately familiar with the client's perspective.

## Our Mission

In partnering with a pharma/biotech company, we strive to:

- ◆ provide the highest quality biostatistical and programming services (first time right quality – no compromise)
- ◆ be timely in completing specified work
- ◆ maintain a very competitive cost



We strive to partner with a client to achieve all of these components leading to a successful collaboration. Our very substantial experience in the pharmaceutical industry - having managed many in-house and outsourced projects - strongly supports this promise. We know what the customer wants and needs in a collaboration. We are very familiar with the issues that can arise when partnering with a CRO, and how to successfully prevent them in the first place, or to effectively deal with them.

## Other Services

- ◆ Exploratory analyses
- ◆ Specialized datasets/analyses, e.g., for specific toxicities and biomarkers
- ◆ CDISC SDTM conversions from legacy datasets
- ◆ Special presentations incorporating statistical results

Please see our website ([www.CRCPharma.com](http://www.CRCPharma.com)) for detailed specification of our services.

## Our Core Services

- ◆ Development of Report and Analysis Plans (statistical and programming specifications)
- ◆ Programming analyses (derived) datasets
- ◆ Generation of Tables, Listings and Graphs for insertion into the Clinical Study Report
- ◆ Generation of Case Report Tabulations (define.pdf or define.xml) for submission to Regulatory Authorities

Note that we have very substantial expertise in early development studies – PK, PD, and proof of concept (POC) studies and are particularly well suited for these studies.

In addition we have considerable experience in later stage studies and submissions.

## CRC Pharma LLC A Clinical Reporting Company

333 Littleton Road, Suite 302  
Parlappany, NJ 07054

Tel: (973) 928 8563  
Fax: (888) 901 7764

E-mail: [ramesh.maddena@crcpharma.com](mailto:ramesh.maddena@crcpharma.com)  
[william.stewart@crcpharma.com](mailto:william.stewart@crcpharma.com)

[www.crcpharma.com](http://www.crcpharma.com)



## 2014 Schedule



**Cambridge, Massachusetts**  
**October 22, 2014**

CDISC: The Good, The Bad, The Standards -  
Best Practices and Lessons Learned  
Novartis Offices, Cambridge, MA  
Registration will open in August.

**Cary, North Carolina**  
**September 29, 2014**

CDISC Tools and Optimization  
SAS Institute Campus, Cary, NC  
Registration will open in July.

Watch [www.PharmaSUG.org](http://www.PharmaSUG.org) for details!

# Consulting Services & Functional Sourcing



Biostatistics | SAS Programming | Data Management | CDISC Conversion

ISO 9000:2008 | CDISC Gold Member

USA | India | Europe



+1 (617) 359-6938 | [info@ephicacy.com](mailto:info@ephicacy.com) | [www.ephicacy.com](http://www.ephicacy.com)

## SAS® AT PHARMASUG 2014

---

We are particularly appreciative to SAS Institute and the many SAS staff for all the help and support they have given to PharmaSUG 2014. Just as SAS is always there for their users, SAS has been here for the organizers of PharmaSUG. A great big “thank you” goes to SAS!

SAS Appreciation Mixer-USS Midway

Tuesday, Jun 3, 6:30PM - 9:30PM

USS-MIDWAY MUSEUM

After a full day of SAS presentations, kick back and relax at the SAS User KICKBACK MIXER. USS Midway - busses start running at 6:00.

Party at 6:30-9:30.

Bus pickup: The location is Gull Street which is on the 1<sup>st</sup> level of the hotel, or Promenade East.

### SAS Exhibits, SAS Demo, and SAS Publications

Monday, Jun 2	9:00 AM to 11:45 AM & 12:45 PM to 6:00 PM	Demo Theater - Exhibit & Demo Room
Tuesday, Jun 3	9:00 AM to 11:45 AM & 12:45 PM to 6:00 PM	Demo Theater - Exhibit & Demo Room
Wednesday, Jun 4	8:00 AM to 11:00 AM	Demo Theater - Exhibit & Demo Room

The Exhibit & Demo Room features an extensive display of SAS products and services. SAS Staff will be on hand to discuss and demonstrate the latest software and services available.

## At Your Service



[support.sas.com/promise](http://support.sas.com/promise)

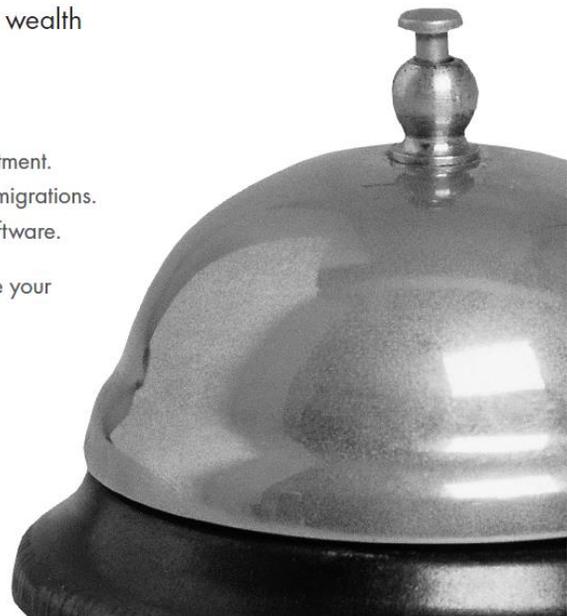
The SAS Customer Loyalty team is your concierge to a wealth of support services and resources.

### How can we assist you?

- Determine which resources can increase the value of your investment.
- Keep you informed about new features, software releases and migrations.
- Make recommendations for getting the most value from your software.

These are just a few examples of what we can do for you. Let us be your guide to all that SAS has to offer.

Contact us at [customerloyalty@sas.com](mailto:customerloyalty@sas.com).



## SAS® PAPERS AND SEMINARS

Below is a list of papers of which SAS developers are one or more of the authors. You can find the abstracts and presentation times for these papers in their respective sections in this conference program.

Presentation	Speaker	Location	Track
Up Your Game with Graph Template Language Layouts	Sanjay Matange	Sapphire Level - L	Data Visualization and Graphics
Negotiation: Getting the Best Out of an Offer	Janet Stuelpner	Exhibit Room Demo Theater	Career Planning
Practically Perfect Presentations	Cynthia Zender	Sapphire Level - P	Hands-On Training
Creating Define-XML version 2 with the SAS® Clinical Standards Toolkit 1.6	Lex Jansen	Sapphire Level - A	Data Standards
An Integrated platform to manage Clinical data, Metadata and Data Standards	Sharon Trevoy   Romain Rutten   Peter Wang	Sapphire Level - A	Data Standards
Clinical Trial Data Transparency: Seeing is Believing	Bill Gibson	Sapphire Level - D	Industry Basics
Risk-Based Monitoring of Clinical Trials Using JMP® Clinical	Kelci Miclaus	Sapphire Level - H	SAS JMP Presentations
Customizing the Kaplan-Meier Survival Plot in PROC LIFETEST in the SAS/STAT® 13.1 Release	Warren Kuhfeld	Sapphire Level - H	Statistics & Pharmacokinetics
When is Validation Valid?	Janet Stuelpner	Sapphire Level - M	Management and Support
A How-To Guide for Extending Controlled Terminology Using SAS Clinical Data Integration	Melissa Martinez	Sapphire Level - A	Data Standards
Creating Multi-Sheet Microsoft Excel Workbooks with SAS®: The Basics and Beyond Part 1	Vince DelGobbo	Sapphire Level - P	Hands-On Training
Modeling Categorical Response Data	Maura Stokes	Sapphire Level - H	Statistics & Pharmacokinetics
Round Trip Ticket – Using the Define.xml file to Send and Receive your Study Specifications	Julie Maddox	Sapphire Level - D	Data Standards

# Discover more.

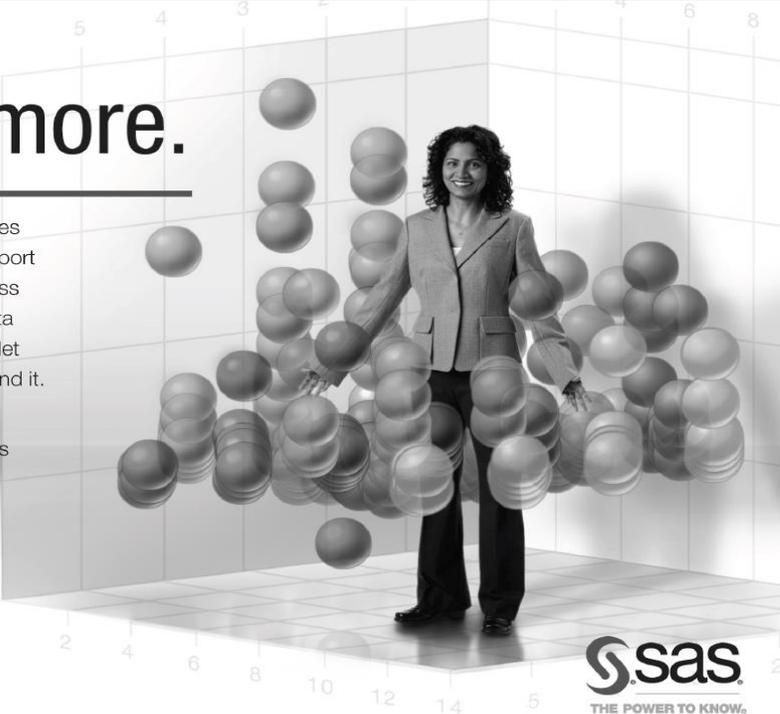
---

Interactive patient profiles and narratives in JMP® Clinical let you monitor and report trial outcomes for individuals and across groups. You see exactly what your data is trying to tell you. Powerful analytics let you explore it. Experience it. Understand it. Then share it with others.

It's data visualization at its best, and it's from SAS.

[jmp.com/clinical](http://jmp.com/clinical)






**Pharma**  
**SUG**  
**CHINA**

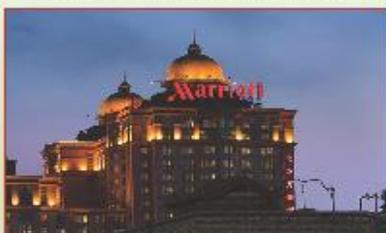


**SEPTEMBER 5-6, 2014**  
**Beijing Marriott Hotel City Wall**

Conference Co-chairs

**Wei J. Chen and Margaret Hung**

info@pharmasugchina.com • www.pharmasugchina.com



***Come for the conference, stay for the culture.***

\*Pre-conference training on September 4th.

## PHARMASUG 2014 TRAINING SEMINARS

Saturday, May 31, 2014

Course Title (click for description)	Instructor(s)	Time
#1 <a href="#">The SDTMIG: Beyond the Modeled Domains and Other Implementation Challenges</a>	<a href="#">Fred Wood</a> & <a href="#">Jerry Salyers</a>	8:00 AM - 12:00 PM
#2 <a href="#">PROC SQL Programming: The Basics and Beyond</a>	<a href="#">Kirk Paul Lafler</a>	8:00 AM - 12:00 PM
#4 <a href="#">CDISC and XML for SAS® Programmers</a>	<a href="#">Lex Jansen</a>	1:00 PM - 5:00 PM
#5 <a href="#">How to Leverage the Power of SAS® Enterprise Guide</a>	<a href="#">Scott Leslie</a>	1:00 PM - 5:00 PM

Sunday, June 1, 2014

Course Title (click for description)	Instructor(s)	Time
#7 <a href="#">How (and Why) to Harmonize Data Collection with CDISC Submission Standards</a>	<a href="#">Steve Kirby</a> & <a href="#">Mario Widel</a>	8:00 AM - 12:00 PM
#8 <a href="#">Cleaning, Validating, and Reshaping Your Data: Taking Advantage of SAS Tools</a>	<a href="#">Art Carpenter</a>	8:00 AM - 12:00 PM
#9 <a href="#">Survival Analysis</a>	<a href="#">Arthur Li</a>	8:00 AM - 12:00 PM
#10 <a href="#">FDA Submission Data Requirements</a>	<a href="#">Dave Izard</a>	1:00 PM - 5:00 PM
#11 <a href="#">Getting Started with Graph Template Language</a>	<a href="#">Sanjay Matange</a>	1:00 PM - 5:00 PM

Wednesday, June 4, 2014

Course Title (click for description)	Instructor(s)	Time
#12 <a href="#">ADaM Datasets: Creation and Applications</a>	<a href="#">Nancy Brucken</a> & <a href="#">Mario Widel</a> & <a href="#">Paul Slagle</a>	1:00 PM - 5:00 PM
#14 <a href="#">Building Dynamic Programs and Applications Using the SAS Macro Language</a>	<a href="#">Art Carpenter</a>	1:00 PM - 5:00 PM

Thursday, June 5, 2014

Course Title (click for description)	Instructor(s)	Time
#15 <a href="#">Advanced Reporting and Analysis Techniques for the SAS® Power User: It's Not Just About the PROCs!</a>	<a href="#">Art Carpenter</a>	8:00 AM - 5:00 PM



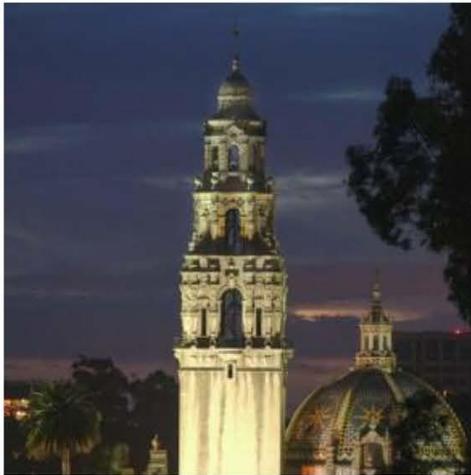
*Things to Do in San Diego, CA*



**Chill Out in La Jolla!!!**



**Tour the USS Midway!!!**



**Check out Balboa Park!!!**



**Check out Coronado Island!!!**



## Have Joined Forces

A Leader in Conducting Full-Service, Global Trials  
The Pioneer of Embedded Outsourcing Solutions

Built for Innovation | Built to Serve You Better

Now Together –  
Transforming Clinical Trials



Learn how PRA & RPS can transform your career at Booth 33

Learn more



Phase I-IV: Biopharmaceutical Drug Development  
Embedded and Full-Service Solutions

*Transforming Clinical Trials through...*  
*Our People - Innovation - Transparency*

# Online Statistics Programs



**STATISTICS**  
TEXAS A&M UNIVERSITY



## **Online Statistics Education from Texas A&M University**

Master of Science in Statistics • Graduate Certificates • Individual Courses

### **Emphasis Areas**

#### **Biostatistics**

*Manage your data from start to finish.*

Biostatistics is one of the best areas to engage in collaborative consulting and research.

#### **Data Mining**

*Enhance your career prospects.*

Whether you are working in banking, insurance, consumer services, or a government agency, you can use business analytics and data mining to make better decisions.

#### **Applied Statistics**

*Prepare for your future.*

Our courses offer you the opportunity to learn how to apply statistics to real problems.

**Visit us at our booth to discuss how we can help you jumpstart your future with online education.**

*Replicating the  
classroom experience with  
Online Learning*



**Stop by our booth to register to win a Kindle Fire**

**[www.stat.tamu.edu/dist](http://www.stat.tamu.edu/dist)**



## Genentech Careers

### Creating Breakthrough Medicines That Make Lives Better

As the first biotechnology company, Genentech is an industry leader with an impressive history of firsts in medicine. But take a closer look and you'll see that we're about so much more. We're about a commitment to patients, science and people. We're about diverse employees coming together to create breakthrough medicines that change the lives of people and help manage disease. We're about encouraging success and providing rewards to those who work here. We're about helping current students and new graduates forge a unique and meaningful career path and ensuring that experienced professionals have the opportunities and resources to excel no matter where they are in their careers. And we're about helping employees create a fulfilling balance between their personal lives and career aspirations.

## 100 Best Companies to Work For Fortune Magazine 2014

*In January 2014, FORTUNE named Genentech among the "100 Best Companies to Work For" for the 16th consecutive year.*

To pick the 100 Best Companies to Work For, FORTUNE partners with the Great Place to Work Institute (GPTWI) to conduct an extensive employee survey of corporate America. This year, 257 firms participated in the survey with more than 252,000 employees at those companies responding to the survey.

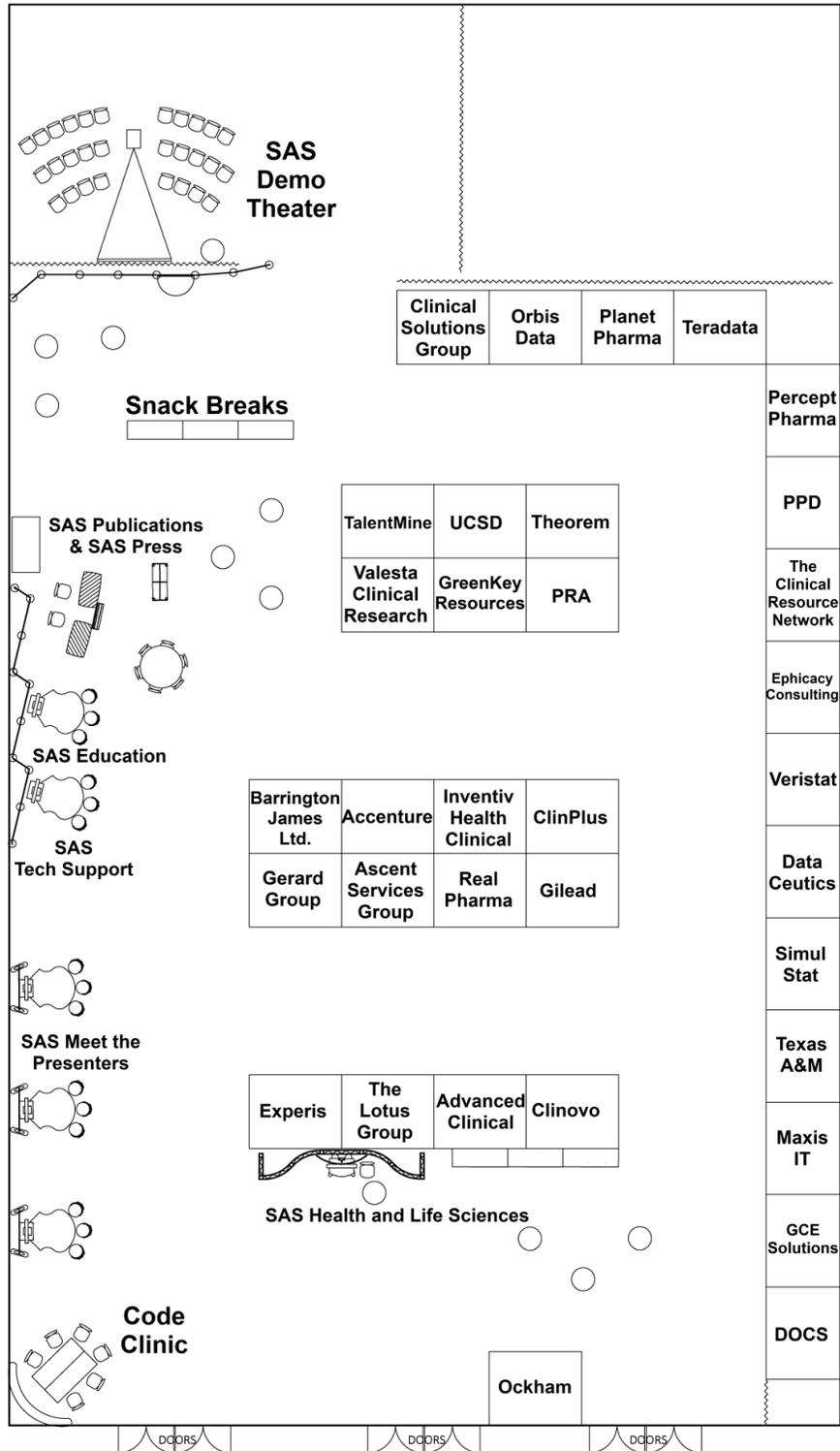
Two-thirds of a company's score is based on the results of GPTWI's Trust Index survey, which is sent to a random sample of employees from each company. The survey asks questions related to employees' attitudes about management's credibility, job satisfaction and camaraderie. The other third is based on responses to GPTWI's Culture Audit, which includes detailed questions about pay and benefit programs and a series of open-ended questions regarding hiring practices, methods of internal communication, training, recognition programs and diversity efforts.

Genentech was also recognized as one of 2014's social media superstars, staying abreast on the latest trends in social media, especially when connecting employees with senior leadership, company news, and each other.

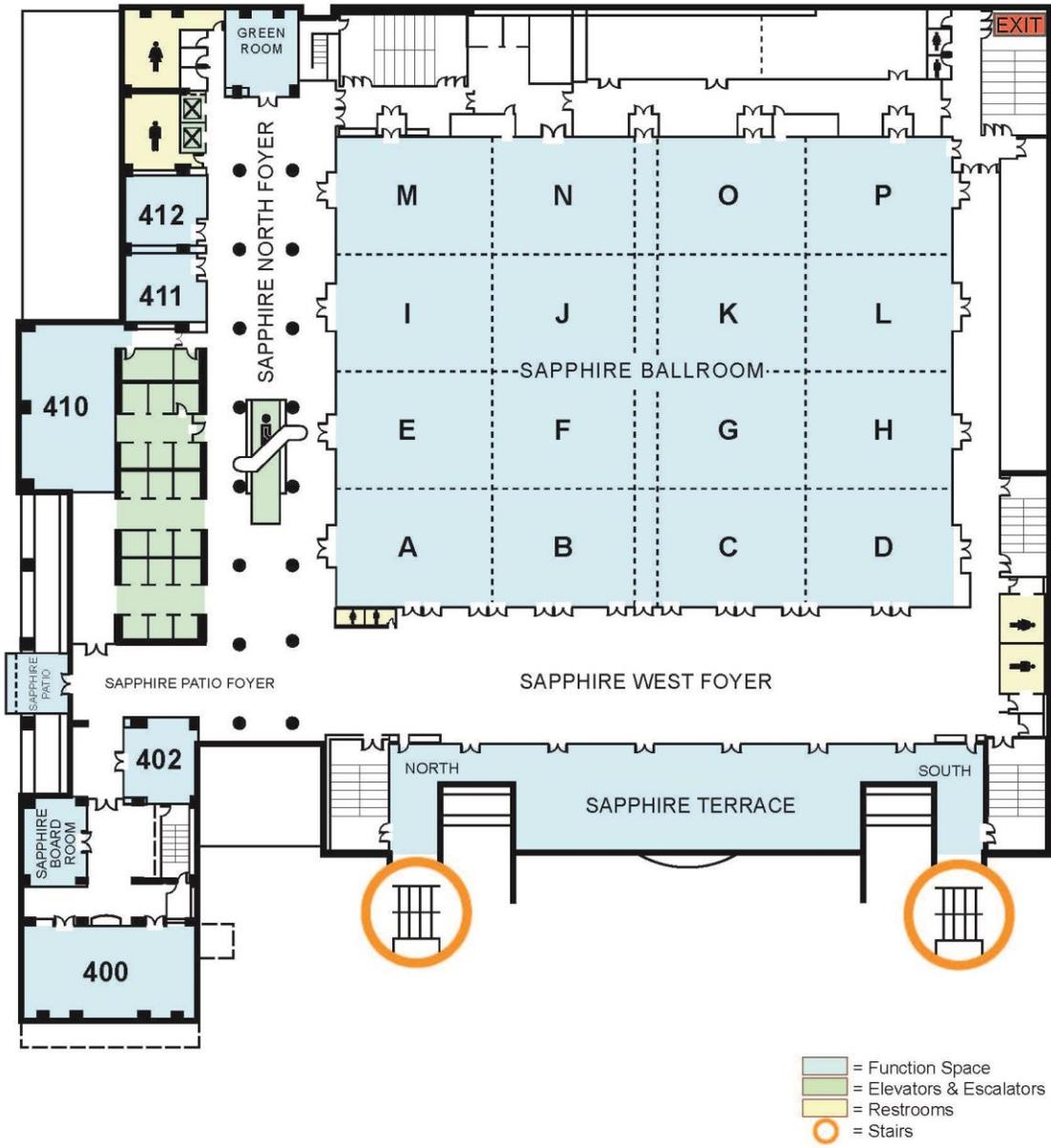


## Exhibit & Demo Room

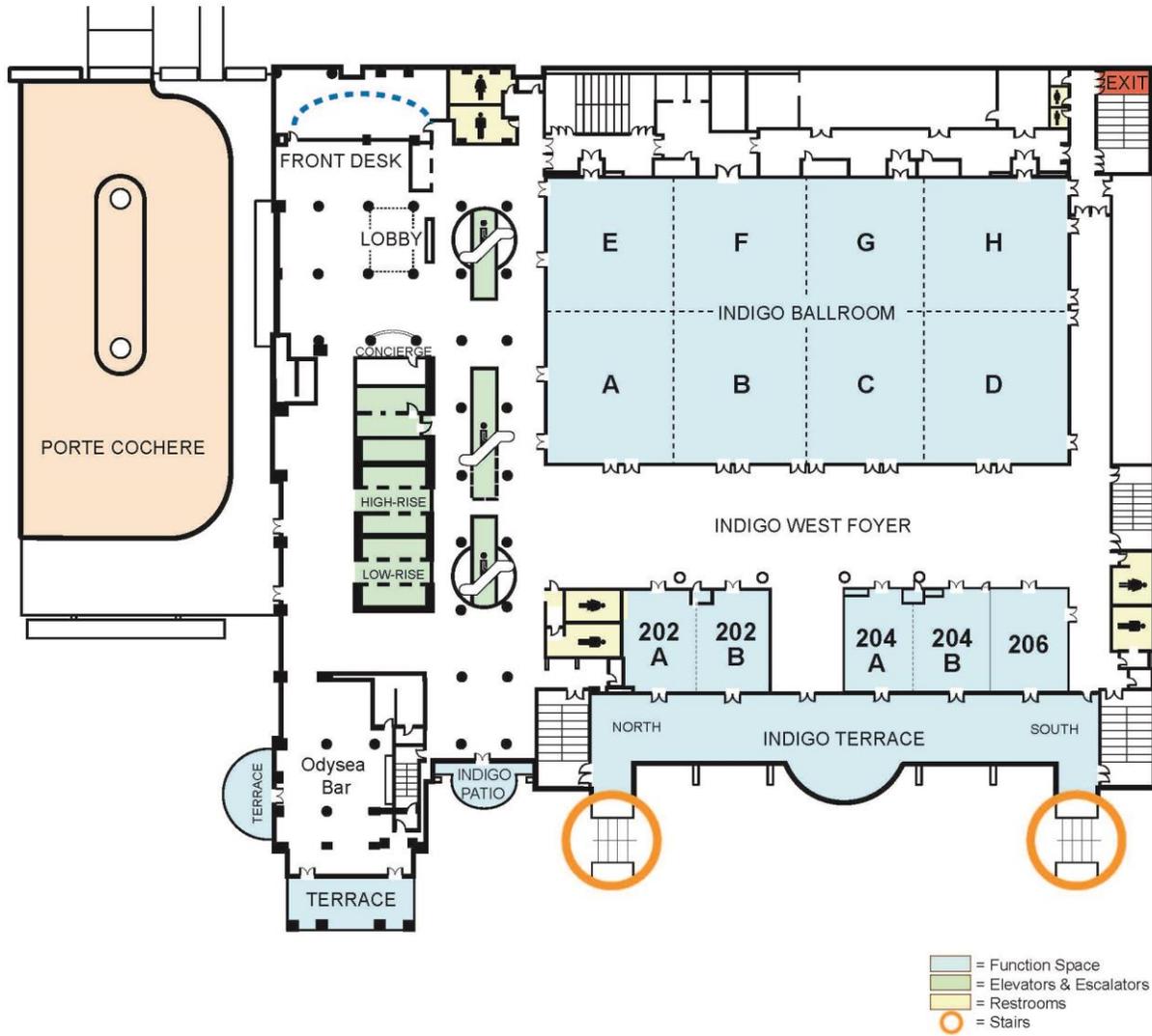
### PharmaSUG 2014 Demo Room and Exhibit Area



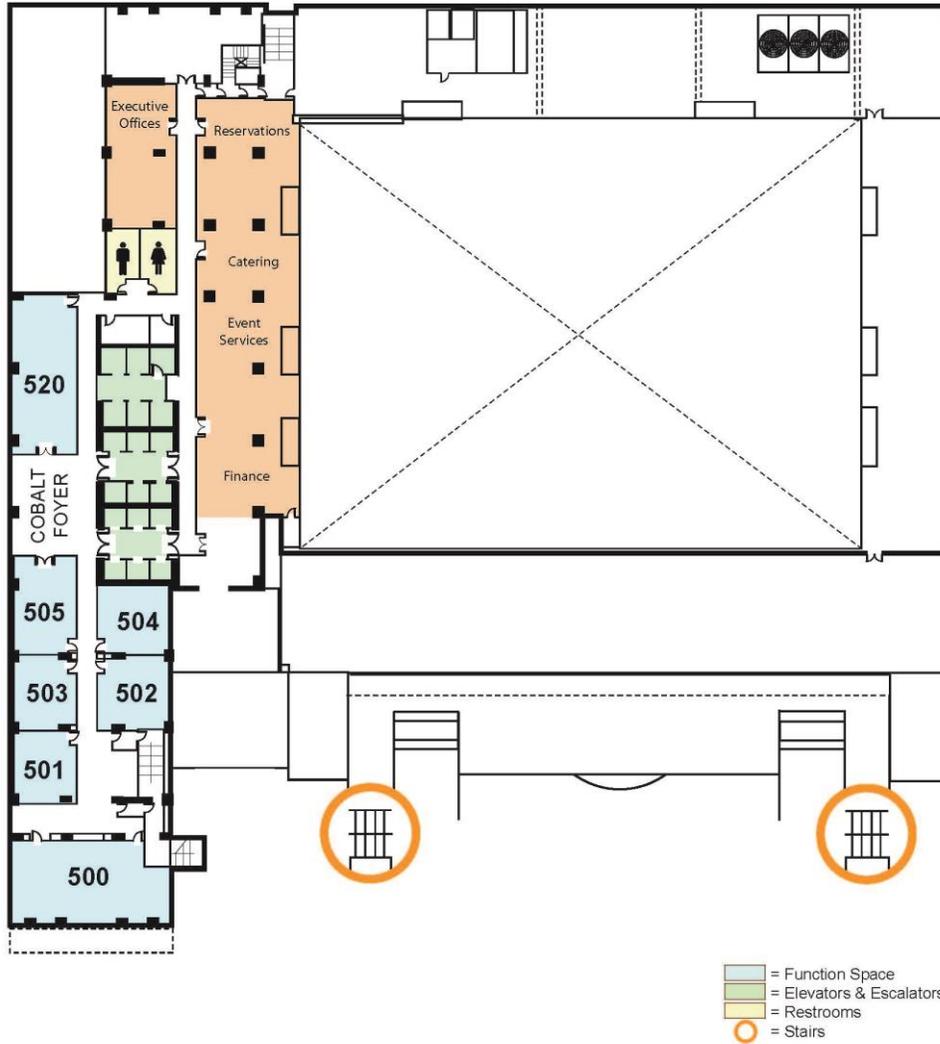
# Sapphire Ballroom



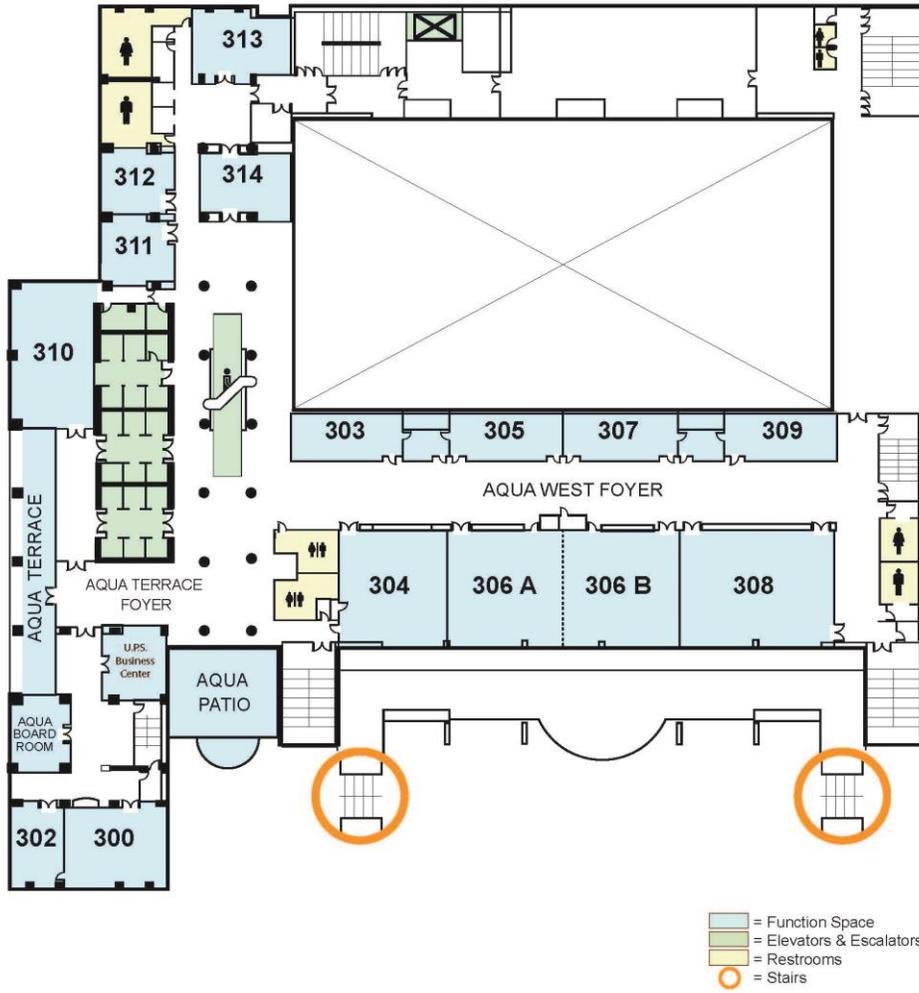
# Indigo Ballroom



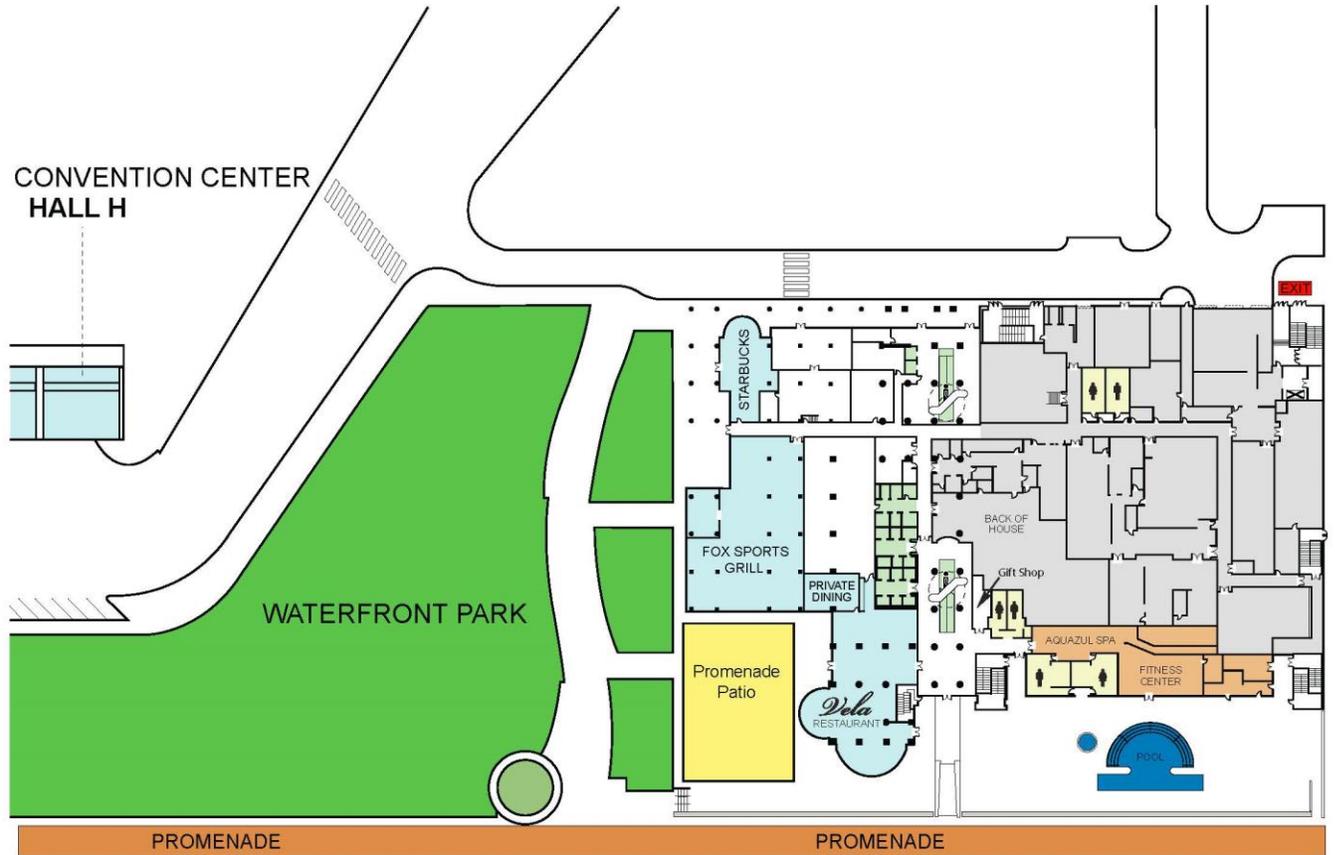
## Cobalt Level



# Aqua Level



# PROMENADE



- Function Space
- Elevators & Escalators
- Restrooms
- Stairs

## Applications Development

---

Paper AD01

Tuesday: 8:00 - 8:20 a.m.

Room: Sapphire I

### **Converting Multiple Plain DOC Files to RTF Files in SAS**

Author(s): William Wu, Xiaoxian Dai, Linda Gau

Creating tables and listings in the Rich Text Format (RTF) has become immensely popular in the pharmaceutical industry. The most common method of creating RTF files in SAS® is to use the PROC REPORT procedure with Output Delivery System (ODS). However, it is not easy to create multiple RTF files with uniform formats. An alternative method is to generate RTF files in two steps, which involves in creation of plain DOC files and conversion of DOC files to RTF files. The first step can be easily accomplished with the combination of PROC REPORT and ODS. Currently, there is no easy way to do the second step. In this paper, we introduce a macro, %doc2rtf, which was developed in SAS 9.3 and can be use to convert multiple plain DOC files in the same folder to RTF files with uniform formats. The DOC files are first read and converted to SAS datasets, which is then used to create RTF files based on Rich Text Format Specification Version 1.9.1 by using data \_NULL\_ step. With the availability of the %doc2rtf macro, multiple RTF files with uniform formats can be easily generated from plain DOC files. Therefore, it saves plenty of time in creating multiple RTF files with uniform formats in SAS.

---

Paper AD02

Tuesday: 10:15 - 11:05 a.m.

Room: Sapphire I

### **Creating Define.xml v2 Using SAS for FDA Submissions**

Author(s): Qinghua Chen

When submitting clinical data to the Food and Drug Administration (FDA), besides the usual trials results, we need to submit the information that helps the FDA to understand the data. The FDA has required the CDISC Case Report Tabulation Data Definition Specification (Define-XML), which is based on the CDISC Operational Data Model (ODM), for submissions using Study Data Tabulation Model (SDTM). Electronic submission to the FDA is therefore a process of following the guidelines from CDISC and FDA. This paper illustrates how to create an FDA guidance compliant define.xml v2 from metadata by using SAS®.

---

Paper AD03

Tuesday: 9:00 - 9:50 a.m.

Room: Sapphire I

### **Table Lookup Techniques: From the Basics to the Innovative**

Author(s): Art Carpenter

One of the more commonly needed operations within SAS® programming is to determine the value of one variable based on the value of another. A series of techniques and tools have evolved over the years to make the matching of these values go faster, smoother, and easier. A majority of these techniques require operations such as sorting, searching, and comparing. As it turns out, these types of techniques are some of the more computationally intensive, and consequently an understanding of the operations involved and a careful selection of the specific technique can often save the user a substantial amount of computing resources. Many of the more advanced techniques can require substantially fewer resources. It is incumbent on the user to have a broad understanding of the issues involved and a more detailed understanding of the solutions available. Even if you do not currently have a BIG data problem, you should at the very least have a basic knowledge of the kinds of techniques that are available for your use.

Paper AD04

Monday: 2:15 - 3:05 p.m.

Room: Sapphire I

### **Deploying a User-Friendly SAS Grid on Microsoft Windows**

Author(s): Houliang Li

Your company's chronically overloaded SAS environment, adversely impacted user community, and the resultant lackluster productivity have finally convinced your upper management that it is time to upgrade to SAS grid to eliminate all the resource problems once and for all. But after the contract is signed and implementation begins, you as the SAS administrator suddenly realize that your company-wide standard mode of SAS operations, i.e., using the traditional SAS Display Manager on a server machine, runs counter to the expectation of SAS grid - your users are now supposed to switch to SAS Enterprise Guide on a PC. This is utterly unacceptable to the user community because almost everything has to change in a big way. If you like to play a hero in your little world, this is your opportunity. There are a number of things you can do to make the transition to SAS grid as smooth and painless as possible, and your users get to keep their favorite SAS Display Manager.

---

Paper AD05

Monday: 11:15 - 12:05 p.m.

Room: Sapphire I

### **The SAS® LOCALE Option Win an international passport for your SAS code**

Author(s): Xiaojin Qin, Peter Eberhardt

Want to win an international passport for a round-the-world journey? Start with your SAS code! SAS internationalization is the step that generalizes your project to be language independent. This allows you to write code once, and then have it run in different cultural environments with the appropriate cultural interpretation; for example, write your code so it uses formats that are familiar to you in China, yet when your French clients run the code it will use formats familiar to them. LOCALE, one of the internationalization features of SAS, can help your code be a world traveler. This paper will show the use of LOCALE in our SAS sessions as a major part of making the round-the-world journey a reality. You will see how a minor change can make a major difference.

---

Paper AD06

Tuesday: 8:30 - 8:50 a.m.

Room: Sapphire I

### **Automatic Access Control via SAS: More Efficient and Smart**

Author(s): Koretaka Yuichi, Fujiwara Masakazu, Yoshida Yuki, Kitanishi Yoshitake

Clinical projects have a wide variety of information. Usually, they are stored at in-house system all together. Some information is very confidential. Since they include confidential information (e.g. clinical summary report, study protocol, clinical data and analysis results etc.), the access rights must be given to only appropriate persons. On the other hand, the other information might be not so confidential (e.g. analysis plan and program code etc.). In that case, such information should be shared so that the knowledge can be effectively used in the department. So, we decided to control access rights depending on the property of information per folders. In addition, our folder structure is very complicated, because it is created in reference to Study Data Specifications that FDA released in 2012 [1]. Therefore, setting access rights manually is not only inefficient but also might cause human error. For these reasons, we developed the SAS programs for setting access rights automatically to each folder. SAS has linkage functions with Excel and Command Prompt and can handle datasets freely more than other software. By making use of these features, our programs allowed IT support team (i.e. administrator) who is not familiar with SAS to manage access rights easily. All they have to do is to input the names and roles (statistician, programmer, or data manager etc.) of users and click an execution button on Excel. The purpose of this paper is to describe our information management paradigm with SAS programs and provide some important snippets of SAS code.

**Techniques for Managing (Large!) Batches of Statistical Output**

Author(s): Brian Fairfield-Carter, Tracy Sherman

When delivering batches of statistical output to a client, some fairly basic conditions need to be met: output must be current (generated without error from current analysis data), validation status must be current, and the collection of output must comprise what the client is expecting to receive (no missing files or extra files). While these conditions are fairly easy to enforce with small batches of output, they can be surprisingly difficult to meet with larger batches: consider for instance the practicality of checking that each of 350 entries in a programming plan has a corresponding output file (or conversely that each file has a corresponding entry). Techniques for managing a small delivery tend to be inadequate for large deliveries, and as the size and complexity of the delivery increases, the need for automation in meeting/confirming these conditions also increases. Since large batches of output are difficult to navigate, clients will often (and sometimes without advance warning) request enhancements such as a hyperlinked Table of Contents (TOC) and/or bookmarked collated file. These can be prohibitively time-consuming to create manually, but automated methods may involve technology that is unfamiliar to most SAS programmers. This paper discusses problems in managing large batches of statistical output, and offers practical and automated techniques for handling the various steps in assembling a delivery, including Batch-updates to production and QC output; Word/RTF document comparisons; Checking for agreement between programming plan and output; and RTF, PDF, & Excel File collation and hyperlinked TOC generation

**Challenges in Moving to a Multi-tier Server Platform**

Author(s): Carey Smoak

Moving from a simple server to a multi-tier server environment poses many challenges. The multi-tier server environment in this paper includes a physical application (SAS® version 9.3) server and a virtual metadata server. No mid-tier (web) server is included in this configuration. The SAS Management Console 9.3 is used to administer the metadata (users, groups, roles and permissions). Two types of clients (SAS Enterprise Guide and the SAS Add-in to Microsoft Office) are used to perform various tasks using the data on the application server. Challenges include: (1) validating SAS on a multi-tier system, (2) understanding and administering metadata using SAS Management Console, and (3) training users on Enterprise Guide and the Add-in to Microsoft Office. The business need for moving to this type of environment was to allow non-SAS users to have read-only access to SAS datasets to monitor and query clinical trial data. This can be accomplished through the Add-in to Microsoft Office. For example, Microsoft Excel can be used to open SAS datasets (read-only) and query them. Furthermore, stored processes (SAS programs) can be written in Enterprise Guide and made available to non-SAS users to run monitoring reports in Excel.

---

Paper AD09

Monday: 10:45 - 11:05 a.m.

Room: Sapphire I

### **Macro Programming via Parameter Look-Up Tables**

Author(s): Joseph Hinson

SAS Macros provide a unique means of creating one-size-fits-all programs and many large institutions rely on secure non-editable macros for generating standardized clinical reports. In such situations, programmers are shielded from the standard macros and are instead simply made to manipulate an external source of parameters in order to produce customized reports. This approach can further be extended by organizing report macros into domain-neutral classes based only on the structural layout of the reports, and another set based on analysis types, such that a call to a parameter look-up table would provide the appropriate combination of structural and analysis class macros pertinent to a particular report. With such an approach, just a dozen macro classes can rely on a large set of parameters in a look-up table to generate hundreds of different customized report across domains and therapeutic areas. The present paper aims to offer the concepts behind such an approach with the hope that application developers and programmers can gain from them innovative ideas for efficient macro programming for the generation of clinical reports.

---

Paper AD10

Tuesday: 11:15 - 11:35 a.m.

Room: Sapphire I

### **DefKit: A Micro Framework for CDISC Define-XML Application Development**

Author(s): Lei Zhang

CDISC Define-XML specification provides an extended ODM model to describe the clinical data and statistical analyses submitted for FDA review. However, because of the inherent complexity, creating define.xml files from clinical studies poses many challenges to SAS programmers. In this paper, we introduce a micro-framework called DefKit to facilitate the development of CDISC/XML applications. DefKit is designed as a rapid development framework, and is built specifically for SAS programmers who want to have a simple way to create CDISC/XML applications, such as define.xml generators. The toolkit that the DefKit framework provides consists of a very small set of macros and user defined functions (UDFs), and employs hash objects for fast data retrieval and XML generation. This paper first introduces the organization and usages of the DefKit framework components, and then provides examples to illustrate how to generate define.xml elements with DefKit. Apart from explaining how to implement these elements, it demonstrates how DefKit's unique features make it easy to develop and maintain CDISC/XML applications .

---

Paper AD11

Tuesday: 11:45 - 12:05 p.m.

Room: Sapphire I

### **Dealing with Missing Data in Clinical Trials**

Author(s): Lucheng Shao

It seems inevitable to encounter missing data in clinical trials no matter how perfect the study was designated and how carefully Clinical Research Associates collected the data. However, having missing values in our original clinical database is not the end of the world for our SAS programmers. The focus of this paper will be showing you how to deal with missing data in clinical trials, especially how to improve the way of representing missing data so that the missing information can be taken advantage of in improving reports. This paper does not cover missing data mechanisms or imputation methods. It is intended for readers who are similar with SAS base but not with different types of missing data.

---

Paper AD12

Tuesday: 1:45 - 2:05 p.m.

Room: Sapphire I

### **MACUMBA - The Log Belongs to the Program**

Author(s): Michael Weiss

MACUMBA is an in-house-developed application for SAS® programming. It combines interactive development features of PC-SAS, the possibility of a client-server environment and unique state-of-the-art features that were always missing. This presentation shows a feature of the MACUMBA application that aligns the generated log file to the SAS program code, it belongs to. In this way all relevant log lines (ERROR's, WARNING's and others) are directly displayed at the source code line, they occurred. This alignment provides an easy way to find the source of a problem and ensures that no important log line is overseen anymore. Additionally for this presentation a small tool was created that provides most of the presented features and is freely available.

---

Paper AD13

Tuesday: 2:15 - 3:05 p.m.

Room: Sapphire I

### **Let SAS Improve Your CDISC Data Quality**

Author(s): Wayne Zhong

How often is "but this data passed OpenCDISC" used as a defense for poor quality data? Adoption of data standards benefits reviewers and analysts by keeping data structures consistent. Ensuring data integrity however remains a manual labor with edit checks or eyeballing. This paper presents a data quality control application for CDISC data: using the metadata of CDISC datasets to trigger SAS code modules, running relevant checks for each dataset and presenting issues for review. This tool can be run with the availability of any CDISC data, not waiting for late statistical analysis to report strange results and later investigations to identify data issues, making this application an important component of a good CDISC data creation process.

---

Paper AD14

Monday: 9:00 - 9:20 a.m.

Room: Sapphire I

### **Add Dropdowns and Redirects to Your SAS Output**

Author(s): Barbara Ross

Outputting attractive and interactive reports via SAS onto the web doesn't have to be complicated. For most internal purposes, a webpage directory that links to static reports that update regularly is enough. This paper introduces simple ways to add dropdown menus and redirects within your SAS output. All webpages are created using Base SAS and website directories are managed using an open source FTP client for Windows (WinSCP in this case). Techniques used include: JavaScript to select item on click, PHP to collect form entries, and HTML for hyperlinks. The SQL procedure and DATA step are used for data transformations.

---

Paper AD17

Tuesday: 3:30 - 4:20 p.m.

Room: Sapphire I

### **The New Tradition: SAS® Clinical Data Integration**

Author(s): Vincent Amoruccio, Janet Stuelpner

Base SAS® programming has been around for a very long time. And over that time, there have been many changes. New and enhanced procedures, new features, new functions and even operating systems have been added. Over time, there have been many windows and wizards that help to more easily generate code that can be used in programs. Through it all, programmers always come back to their SAS roots, the basic programs with which they started. But, as we move into the future, is this the best use of time, to sit and manually code everything? Or can we take advantage of the new tools and solutions that generate code and use metadata to describe data, validate output and document exactly what the programmer has done. This paper will show you how we can change the current process using the graphical user interface of SAS Clinical Data Integration to integrate data from disparate data sources and transform that data into industry standards in a methodical, repeatable, more automated fashion.

---

**Don't Type, Talk SAS**

Author(s): Vikas Gaddu

Are you a lazy but smart programmer? Then you will like my paper!! I have seen that SAS programs over last decade are starting to become more and more standardized. We have standard headers, standard init files, standard macros and a standard way to output a Report. In turn we are required to remember lot of standards and their documentation. What if we can talk to SAS and these standard structures would automatically appear in your code? This is now possible with Windows Speech Recognition Macros. Speech recognition has improved tremendously over the last decade. Take it from a person with accent. We have lot of commercially available voice recognition software that can interpret human speech and convert it into text or listen for a keyword and perform bunch of Task. In this paper we are going to discuss freely available software in all windows PC called as Windows Speech Recognition and see how mundane task like writing standard header, standard macros and comments can be easily done using windows speech recognition. With combination of XML and Javascript we can build some very powerful Windows speech recognition macros. This paper assumes that audience is familiar with Voice Recognition, XML and Javascript concepts.

---

**Tips for Creating SAS-Based Applications for Oracle Clinical**

Author(s): Kunal Agnihotri, Ken Borowiak

Oracle Clinical (OC) is database management system used to support clinical trials processes. There is an enormous amount of data and metadata stored in OC beyond the data points collected on a case report form, which can be used as the basis for developing applications to support a variety of activities. This paper offers some tips for creating SAS macros and standard programs against OC. Topics include using PROC SQL Pass-Through facility, resolving macro variables, handling of dates, query optimization and regular expression support. Some useful queries against OC are used to demonstrate each topic.

---

**Managing bulk SAS job submissions with post execution analysis and notification.**

Author(s): Bruce Kayton

Running all the programs in a study analysis with all their dependencies can be a time consuming or inefficient process if done individually or sequentially. This utility uses a driving process to define dependencies, enables programs to run efficiently in quick succession or often simultaneously to best utilize available resources. Upon the completion of all programs a submission summary is generated with run times and log analyses highlighting any warning or error issues. The submitter is notified via email about job completion with a hyperlink to the log analysis in spreadsheet format containing individual program details and hyperlinks to logs and errors. Benefits: Hands off submission, gives users the ability to focus on other tasks. Efficient use of resources allowing programs that are not dependent on each other to be run simultaneously. Instant awareness of relevant issues upon job completion that can also be tied in to an automated notification to responsible parties. Quick access to program logs and error conditions. Summary information gives management a high level view of the status of an analysis. Quantified summary detail to determine time required to consistently run a full analysis. Standardized error and warning evaluation.

---

**Clinical Trial Data Integration: The Strategy, Benefits, and Logistics of Integrating Across a Compound**

Author(s): Keith Hibbetts, Natalie Reynolds

From the 2012 FDA/PhUSE Computational Science Symposium Working Group 3 Vision Statement: "Data integration has always been a challenge both for industry and the agency." One approach to address the challenges of integrating and converting data across studies is to build a clinical trial integrated database (IDB). This paper explores the strategy of creating and maintaining IDBs across studies and throughout the life cycle of a compound, the benefits IDBs provide, and an effective and efficient metadata-driven system to deal with the logistics of delivering a portfolio of IDBs. First, we'll explore the principles behind the strategy of clinical trial data integration. These principles drive decisions of when to create and update IDBs for a compound. Next, we'll discuss how IDBs provide cost and time savings in supporting many objectives including regulatory submissions, answering customer questions and data mining. Finally, we'll examine the logistics of delivering on that portfolio by examining the meta-data driven system that allows efficient delivery across different compounds. Meta-data is commonly used to define standards across the industry, but this system illustrates the benefits of using metadata to define data transformations as well. By utilizing meta-data concepts, the same sets of tools can be utilized regardless of the standards used in the input studies. Readers of this paper will leave with a solid foundation of the strategy, benefits, and logistics of clinical trial data integration.

**Automation of Appending TLFs**

Author(s): Bill Coar

SAS reports presenting data from clinical trials are typically in the form of summary tables, listings, and figures (TLFs). Whether a small or large set of reports, a single document that appends output files containing bookmarks and a hyperlinked table of contents may better facilitate review and simplify document management. Although numerous techniques exist to append reports, there is much room for improvement when creating a single document composed of multiple TLFs as well as a table of contents directly from SAS. The proposed process takes advantage of external data used to track SAS programming, ODS Document, and Proc Document. Generation of the concatenated report requires two primary functions: identification of TLFs to be appended, and then the actual appending. Many companies already have infrastructure for tracking generation and testing individual TLFs produced using SAS. Provided SAS can communicate with the tracking component of the existing system, a list of outputs with the desired sectioning and sorting can be obtained. The proposed concatenation process is an extension of an application previously presented at PharmSUG 2013. It requires the use of ODS Document to create individual item stores during the initial creation of each TLF, and the use of Proc Document to manipulate and replay such item stores into a single, well-structured document with a hyperlinked table of contents and bookmarks. An example showing how to prepare such a report will be presented using SAS 9.3 in a Windows environment.

Paper AD23

Monday: 1:15 - 2:05 p.m.

Room: Sapphire I

### **Managing the Organization of SAS® Format and Macro Code Libraries in Complex Environments on Multiple SAS® Platforms**

Author(s): Roger Muller

The capabilities of SAS® have been extended by the use of macros and custom formats. SAS macro code libraries and custom format libraries can be stored in various locations, some of which may or may not always be easily and efficiently accessed from other operating environments. Code can be in various states of development ranging from global organization-wide approved libraries to very elementary "just-getting-started" code. Formalized yet flexible file structures for storing code are needed. SAS user environments range from standalone systems such as PC SAS or SAS on a server/mainframe to much more complex installations using multiple platforms. Strictest attention must be paid to (1) file location for macros and formats and (2) management of the lack of cross-platform portability of formats. Macros are relatively easy to run from their native locations. This paper covers methods of doing this with emphasis on: (a) the option "sasautos" to define the location and the search order for identifying macros being called, and (b) even more importantly-the little-known SAS option MAUTOLOCDISPLAY to identify the location of the macro actually called in the saslog. Format libraries are more difficult to manage and cannot be created and run in a different operating system than that in which they were created. This paper will discuss the export, copying and importing of format libraries to provide cross-platform capability. A SAS macro used to identify the source of a format being used will be presented.

---

Paper AD24

Wednesday: 10:15 - 11:05 a.m.

Room: Sapphire I

### **Moving to SAS Drug Development 4**

Author(s): Magnus Mengelbier

Life Science organisations have a long investment into business processes, standards and conventions that make it difficult to simply turn to a new generation of analysis environments. SAS Drug Development 4 integrates many key features found in current analysis environments that are spread across several applications and systems, which require to be monitored and managed accordingly. The paper considers a set of SAS programs and how the SAS Drug Development repository, workspace and workflow features support a common business process with all of the associated tools and utilities. The result is a short list of points to consider and some tricks for moving a business process from PC SAS or SAS server environments to the new release of SAS Drug Development.

---

Paper AD26

Wednesday: 9:00 - 9:20 a.m.

Room: Sapphire I

### **Validating Analysis Data Set without Double Programming - An Alternative Way to Validate the Analysis Data Set**

Author(s): Linfeng Xu

This paper will demonstrate an alternative way to validate the analysis data set without double programming. The common practice for the most critical level validation of analysis data set in pharmaceutical industry is double programming, that is, the source programmer generates an analysis data set (i.e. production data set) and another programmer (Reviewer) uses same specifications to create a QC data set. Reviewer then conducts PROC COMPARE to see whether or not there is any discrepancy between the two data sets. This paper will introduce ALTVAl macro as an alternative to double programming in validating analysis data set. The ALTVAl macro consists of two parts: 1) compare common variables between raw and analysis data set and use PROC COMPARE to check for discrepancies. (common variable check) 2) using unique merging variable(s), merge the raw and analysis data set to create a new data set, conduct cross-frequency check between input variables (cross check); 3) Based upon step two created big combined data set, reviewer can write simple codes to report the cases which could not meet the variable derivation rules in specification since we have input variable and derived variable in same big combined data set (the logic check of variable specification). This paper will discuss the conceptual design, macro parameters, and prerequisites using this new approach. It will also discuss what types of analysis data sets are suitable for this new validation approach.

**Expediting Access to Critical Pathology Data**

Author(s): Rebecca Ottesen

Abstracting information from pathology notes is often quite cumbersome. Clinical Research Associates and Tumor Registrars typically have to read through all of the diagnosis information and manually enter the data into a database for fields such as tumor size, lymph nodes, and staging. This can lead to data that is subject to interpretation and data entry errors, in addition to the workload burden. In this presentation, we will demonstrate approaches to simplifying data abstraction from pathology reports using various SAS® programming techniques. First, we consider the use of semi-coded College of American Pathologists (CAP) synoptic worksheet data based on a check list filled out by pathologists when reviewing a surgical specimen. We also evaluate the approach of performing string searches on the unstructured pathology dictation using functions such as INDEX() and SUBSTR(). Finally, we demonstrate a validated approach of identifying data from the pathology diagnosis using SAS Text Miner. With all of these tools at hand, it is much easier to meet research needs and to analyze pathology data efficiently.

---

## **Beyond the Basics**

---

Paper BB01

Tuesday: 4:30 - 4:50 p.m.

Room: Sapphire E

### **Indexing: A Powerful Technique for Improving Efficiency**

Author(s): Arun Raj Vidhyadharan, Sunil Jairath

The primary goal of any programmer is "to get the desired output". Once we have a plan to achieve this primary goal, there should ideally be a secondary goal, which is "to get the desired output, efficiently". Most programmers are successful in achieving the primary goal while at least some don't pay much attention or don't feel the necessity of the secondary goal. This paper focuses on a technique called indexing in SAS that could drastically improve the performance of SAS programs that access small subsets of observations from large SAS data sets. This paper covers creation of simple and composite indexes, determining ideal candidates for index key variables, understanding when to use index and how to generate index usage messages.

---

Paper BB02

Monday: 2:15 - 3:05 p.m.

Room: Sapphire E

### **Understanding and Applying the Logic of the DOW-Loop**

Author(s): Arthur Li

The DOW-loop is not official terminology that one can find in SAS® documentation, but it was well-known and widely-used among experienced SAS programmers. The DOW-loop was developed over a decade ago by a few SAS gurus, including Don Henderson, Paul Dorfman, and Ian Whitlock. A common construction of the DOW-loop consists of a DO-UNTIL loop with a SET and a BY statement within the loop. This construction isolates actions that are performed before and after the loop from the action within the loop which results in eliminating the need for retaining or resetting the newly-created variables to missing in the DATA step. In this talk, in addition to understanding the DOW-loop construction, we will review how to apply the DOW-loop to various applications.

---

Paper BB03

Monday: 3:30 - 3:50 p.m.

Room: Sapphire E

### **That SAS®sy Lab Data**

Author(s): Amie Bissonett

Working with laboratory data can be a daunting task. Data can come from central labs, local labs, or both on the same study. Receiving lab data from different labs can result in values within a lab test to have differing units and normal ranges, which need to be reconciled prior to analyzing and displaying results so that values are accurate and consistent. CDISC standards bring additional issues to the table that need to be considered while programming the SDTM and ADaM data sets. This paper displays some data issues to look for and how to handle them, as well as how to handle programming issues for SDTM, ADaM, and CTCAE grading.

Paper BB04

Monday: 4:00 - 4:50 p.m.

Room: Sapphire E

**D is for Dynamic, Putting Dynamic Back into CDISC (A simple macro using PROC SQL which auto-formats continuous variables)**

Author(s): Steven C. Black

As CDISC implementation increases, the ability and need for simple dynamic code has also risen. Using PROC SQL and the SAS MACRO language coupled with the proper use of the ADAM, generating dynamic code has become much easier. This paper illustrates one way of auto generating the decimal format needed for tables where continuous data is arranged by parameters. This code has been especially helpful with laboratory table values where the decimal formats change between each parameter. Understanding and using these same techniques and framework, additional dynamic code can be created allowing you to put the D back into CDISC.

---

Paper BB05

Tuesday: 8:00 - 8:50 a.m.

Room: Sapphire E

**I Object: SAS® Does Objects with DS2**

Author(s): Peter Eberhardt, Xue Yao

The DATA step has served SAS® programmers well over the years, and although it is powerful, it has not fundamentally changed. With DS2, SAS has introduced a significant alternative to the DATA step by introducing an object-oriented programming environment. In this paper, we share our experiences with getting started with DS2 and learning to use it to access, manage, and share data in a scalable, threaded, and standards-based way.

---

Paper BB07

Monday: 10:15 - 11:05 a.m.

Room: Sapphire E

**Five Ways to Flip-Flop Your Data**

Author(s): Josh Horstman

Data are often stored in highly normalized ("tall and skinny") structures that are not convenient for analysis. The SAS® programmer frequently needs to transform the data to arrange relevant variables together in a single row. Sometimes this is a simple matter of using the TRANSPOSE procedure to flip the values of a single variable into separate variables. However, when there are multiple variables to be transposed to a single row, it may require multiple transpositions to obtain the desired result. This paper describes five different ways to achieve this "flip-flop", explains how each method works, and compares the usefulness of each method in various situations. Emphasis is given to achieving a data-driven solution that minimizes hard-coding based on prior knowledge of the possible values each variable may have and improves maintainability and reusability of the code. The intended audience is novice and intermediate SAS programmers who have a basic understanding of the data step and the TRANSPOSE procedure.

---

Paper BB08

Monday: 5:00 - 5:20 p.m.

Room: Sapphire E

**Express Yourself! Regular Expressions vs SAS Text String Functions**

Author(s): Spencer Childress

SAS® and Perl regular expression functions offer a powerful alternative, and complement, to typical SAS text string functions. Harnessing the power of regular expressions, SAS functions such as PRXMATCH, PRXCHANGE, and CALL PRXSUBSTR not only overlap functionality with functions such as INDEX, TRANWRD, and SUBSTR, they also eclipse them. With the addition of arguments 3 and 4 to such functions as SCAN, COMPRESS, and FIND, some of the regular expression syntax already exists for programmers familiar with SAS 9.2 and later. We look at different methods that solve the same problem, with detailed explanations of how each method works. Problems range from simple searches to identifying common text strings like dates. Programmers should expect an improved grasp of the regular expression and how it can complement their portfolio of code. The techniques presented herein offer a good overview of basic data step text string manipulation appropriate for all levels of SAS capability. While this article targets a clinical computing audience, the techniques apply to a broad range of computing scenarios.

---

Paper BB10

Tuesday: 9:00 - 9:50 a.m.

Room: Sapphire E

### **Using the power of SAS SQL**

Author(s): Jessica Wang

SAS is a flexible language in that the same task can be accomplished in numerous ways. SAS SQL is a powerful tool for data manipulation and query. SAS SQL can make your programs more efficient, simpler and more readable. Topics in the paper will cover: merging multiple tables by different columns and different rules; SQL in-line view; SQL set operation; using dictionary tables; creating empty tables with pre-defined dataset structure; inserting rows into a dataset; assigning a list of macro variables. Some tricks and tips from the author's personal experience as a SAS user will also be shared: e.g. using COMPRESS to save storage size; using SQL options \_METHOD to understand your SQL code better; using NOPRINT to compress unnecessary display in output window. This paper is intended for intermediate to advanced SAS SQL users who already know the basics of SAS/SQL, and want to better exploit the power that SQL offers.

---

Paper BB11

Tuesday: 10:15 - 11:05 a.m.

Room: Sapphire E

### **XML in a SAS and Pharma World**

Author(s): Mike Molter

The introduction of standards to the clinical data life cycle has brought about significant changes to the job of a SAS programmer. Traditionally, SAS programmers who built data sets and statistical output for regulatory submission purposes had little need for technical knowledge beyond SAS programming and maybe some Microsoft or Adobe basics. In today's world, collected data can be made available in an XML format produced by EDC systems (i.e. ODM.xml). Submissions are expected to be accompanied by metadata expressed through an XML extension of ODM (i.e. Define.xml). There's even talk of replacing the traditional version 5 transport files with another XML extension of ODM (i.e. SDS-XML) for submission of domain data. The increased use of XML for transferring data and metadata has its advantages, but also imposes new requirements of a programmer's skill set. This paper serves as an introduction to some of these new skills. In it we'll discuss general XML basics and examine how they are applied in our industry today. We'll also look at tools that allow us to move data from a SAS data set to an XML format and vice versa. This paper is for industry SAS programmers at an intermediate level or above.

---

Paper BB12

Tuesday: 2:00 - 2:20 p.m.

Room: Sapphire E

### **Basics of Macro processing - Q way**

Author(s): Usha Kumar

Macros have been one of the challenging areas of SAS programming. This paper is an attempt to make the fundamentals of macro working easy through the 'Q' way. Yes, I call it the Q way as we will look at the Macro Quoting functions to understand the working of Macros. There are numerous articles on the Macro Quoting functions. The difference here is that I am using these functions to present them from a different perspective; i.e. to understand the stages of Macro processing. When it comes to Macros, we must understand the difference between 2 key terms "Compilation" and "Execution". These are different stages of processing. Knowing this difference will help understand Macros better and interpret/debug the Macro errors faster.

Paper BB13

Tuesday: 11:15 - 12:05 p.m.

Room: Sapphire E

### **Kaplan-Meier Survival Plotting Macro %NEWSURV**

Author(s): Jeffrey Meyers

The research areas of pharmaceuticals and oncology clinical trials greatly depend on time-to-event endpoints such as overall survival and progression-free survival. One of the best graphical displays of these analyses is the Kaplan-Meier curve, which can be simple to generate with the LIFETEST procedure but difficult to customize. Journal articles generally prefer that statistics such as median time-to-event, number of patients, and time-point event-free rate estimates be displayed within the graphic itself, and this was previously difficult to do without an external program such as Microsoft Excel. The macro NEWSURV takes advantage of the Graph Template Language (GTL) that was added with the SG graphics engine to create this level of customizability without the need for backend manipulation. Taking this one step further, the macro was improved to be able to generate a lattice of multiple unique Kaplan-Meier curves for side by side comparisons or condensing figures for publications. The following is a paper describing the functionality of the macro, a description of how the key elements of the macro work, and the actual macro code itself.

---

Paper BB14

Tuesday: 5:00 - 5:20 p.m.

Room: Sapphire E

### **A SAS® Macro Utility to Modify and Validate RTF Outputs for Regional Analyses**

Author(s): Jagan Mohan Achi, Joshua Winters

Clinical Study Reports (CSRs) of clinical trials require the development of the tables, listings, and figures (TLFs) in Rich Text Format (RTF) outputs. The analyses are usually complex and the number of RTF outputs is fairly large depending on the magnitude of the studies. It is not uncommon to see requests to repeat the analyses by region and/or country for multicenter trials to understand the differences in geographical locations. For regional analyses, changes to an RTF output can be but not limited to changing the filename, adding additional information in the title, and modifying the hidden bookmark which is used to ensure that the report refers to the specific TLF from CSRs. Such changes have to be made to the original TLF programs which are time consuming. Therefore, a macro was developed to systematically modify the RTF outputs, save the processed RTF outputs with new names, and validate the original and modified RTF outputs to confirm that the only intended changes were made to the RTF outputs. Keywords: CSR, Regional Analyses, RTF

---

Paper BB15

Tuesday: 3:30 - 4:20 p.m.

Room: Sapphire E

### **Atypical application of PROC SUMMARY**

Author(s): John King

This paper describes using PROC SUMMARY combined with a few other DATA and PROC steps to produce stacked frequency tables of a large or small number of categorical variables. This technique uses a single pass of the analysis data preserving the variable order and the order of the levels of each variable, while providing complete rows and complete columns of zero counts (something from nothing) and BIGN. We will discuss applications of PROC FORMAT options NOTSORTED and MULTILABEL; PROC SUMMARY options CHARTYPES, PRELOADFMT, COMPLETETYPES, DESCENDTYPES, LEVELS, WAYS, ORDER=DATA, and the TYPES statement; the PROC FREQ WEIGHT statement, ZEROS option and ODS table CROSSTABFREQS; a data step view; PROC TRANSPOSE statements ID and IDLABEL; interesting functions FINDC, VNAME, VLABEL, VVALUE, VFORMATD, CATX and CATS all in the context of a simple extensible program that you can use every day.

Paper BB16

Tuesday: 1:15 - 1:35 p.m.

Room: Sapphire E

### **Automated Validation - Really?**

Author(s): Sneha Sarmukadam

In Clinical Programming, data quality takes priority over all other aspects. Since, it is health-related data that is being processed, it becomes crucial that accurate analysis is performed and the results are displayed exactly the way they are in the analysis datasets. There are two ways to validate outputs - Manual method and Automated method. It has been observed that Automated validation has superseded Manual validation lately due to many factors such as efficiency, 100% data check, code reusability, etc. However, even if Automated Validation has successfully checked the data, it does not mean that the output generated is perfect. Automated Validation methods are known to have some limitations [1], but there are some other underlying issues which may not be spotted easily and can be missed out. This paper tries to highlight such errors and oversights that occur during validations at data display level which can lead to serious quality issues.

---

Paper BB17

Tuesday: 2:15 - 3:05 p.m.

Room: Sapphire E

### **Come Out of Your Shell: A Dynamic Approach to Shell Implementation in Table and Listing Programs**

Author(s): Paul Burmenko, Tony Cardozo

In clinical trials, the SAS programmer uses the Statistical Analysis Plan (SAP) provided by the statistician to create Table, Figure and Listing (TFL) programs to support the clinical trial submission. Table and listing shells or mock-ups from the SAP are the SAS programmer's roadmap in generating clinical study reporting outputs. Statisticians or SAS programmers spend many hours to create shells manually in a word processing application, and then a SAS programmer must spend additional time mimicking the shell's text content and formatting in a table or listing SAS program. What if the shell designer's efforts could be harnessed to reduce the programmer's efforts? This paper presents a case study of how shell content and rich text formatting can be automatically read in from a shell document to create a SAS PROC REPORT shell as a starting point for a table or listing SAS program.

---

Paper BB18

Monday: 9:00 - 9:50 a.m.

Room: Sapphire E

### **Risk-Based Approach to Identifying and Selecting Clinical Sites for Sponsor's Preparation for FDA/EMA Inspection**

Author(s): Xiangchen (Bob) Cui

In December 2012, the Center for Drug Evaluation and Research (CDER) issued a draft guidance relating to electronic submissions. Guidance for Industry: Providing Submissions in Electronic Format-Summary Level Clinical Site Data for CDER's Inspection Planning[1] [2] is one in a series of guidance documents intended to assist sponsors making certain regulatory submissions to FDA in electronic format. FDA's Office of Scientific Investigation (OSI) requests the sponsor to submit a clinical dataset that describes and summarizes the characteristics and outcomes of clinical investigation at the level of the individual study site within all NDAs, BLAs, or supplements that contain clinical data submitted to CDER. The OSI has developed and is piloting a risk-based inspection site selection tool to facilitate use of a risk-based approach for the timely identification of clinical investigator sites for on-site inspection by the CDER during the review of marketing applications. The CDER approved two NDAs (hepatitis C and cystic fibrosis) from Vertex Pharmaceuticals Incorporated in 2011 and 2012, respectively. This paper explores the risk-based methodology, which was developed based on these two NDAs, by analyzing summary level clinical site data to identify and select high risk sites to assist the sponsor in preparation for FDA/EMA inspections. The methods were applied retrospectively to a hepatitis C FDA/EMA submission and prospectively to a cystic fibrosis FDA/EMA submission, both of which were very successful. The sharing of hands-on experiences in this paper is intended to assist readers to apply this methodology to prepare cost-effectively for FDA/EMA inspections through the risk-based approach.

Paper BB19

Monday: 11:15 - 12:05 p.m.

Room: Sapphire E

**Defensive Programming, Avoiding the Big Mistakes**

Author(s): David Franklin

Typically, when you build a small garden shed you should first do some planning, then do construction, and finally look it over and check what you have done and that it is going to do as intended. A similar process is used for writing programs - we should first do some planning, construct the program, and check it over to see that it producing what was asked. This paper takes a brief and lighthearted look at each of these stages, providing a few tips for avoiding some of the many pitfalls and gives a few pieces of SAS code that are useful in development of your program, hopefully avoiding having to say that dreaded phrase, "delete and start again!".

---

Paper BB20

Monday: 1:15 - 2:05 p.m.

Room: Sapphire E

**Efficient SQL for Pharma... and Other Industries**

Author(s): Chris Olinger

Proc SQL has been used for many years by many different types of SAS programmers. This paper delves into how Proc SQL processes internally, and how the user can exploit this information to improve performance with larger data sets. We will go over common tricks such as use of `_method`, as well as talk about views, updates, and the implications of sort and IO "hints" you can specify to Proc SQL to perform better. This talk is intended for persons already familiar with Proc SQL, and SAS configuration options.

## Coders Corner

---

Paper CC01

Monday: 9:00 - 9:10 a.m.

Room: Sapphire D

### **How to Keep Multiple Formats in One Variable after Transpose**

Author(s): Mindy Wang

In clinical trials and many other research fields, proc transpose are used very often. When many variables with their individual formats are transposed into one variable, we lose the formats. We can do a series of if then statements to put the formats back in. However, when the variables involved are too many, the above method can be very tedious. This paper illustrates how to extract formats from dictionary.columns or sashelp.vcolumn, and then use PUTN function to assign the formats at run time and make the task much easier. In addition, it is much easier to apply the same method to other projects without a lot of hard coding in the SAS program. Efficiency is largely increased with this method.

---

Paper CC02

Monday: 9:15 - 9:25 a.m.

Room: Sapphire D

### **Let Hash SUMINC Count For You**

Author(s): Joseph Hinson

Counting of events is inevitable in clinical programming and is easily accomplished with SAS® procedures such as FREQ, MEANS, SUMMARY, TABULATE, SQL or even by simple DATA step statements. In certain situations where counting is a bit intricate involving data partitioning and classifying, another convenient and efficient way is by hash programming. Within the DATA step, hash objects make data directly available to summary variables without the need to first save the data into a dataset. A little-known hash utility, SUMINC, was introduced with SAS® version 9.2. SUMINC is an argument tag for the hash object declaration statement in which it designates a numeric variable for counting. Together with other new SAS 9.2 hash methods REF() and SUM(), counting of items becomes feasible in a key-based manner thus allowing any pattern of counting to be easily accomplished and directly available for summary reporting, as demonstrated in this paper for clinical trial protocol deviation analysis. SUMINC is also useful for NWAY summarization of data, but unlike the SUMMARY procedure or PROC SQL with GROUP BY, NWAY summarization with SUMINC can be combined with DATA step logic within the same DATA step.

---

Paper CC03

Monday: 1:15 - 1:25 p.m.

Room: Sapphire D

### **Quotes within Quotes: When Single (') and Double (") Quotes are not Enough**

Author(s): Art Carpenter

Although it does not happen every day, it is not unusual to need to place a quoted string within another quoted string. Fortunately SAS® recognizes both single and double quote marks and either can be used to within the other. This gives us the ability to have two deep quoting. There are situations, however where two kinds of quotes are not enough. Sometimes we need a third layer or more commonly we need to use a macro variable within the layers of quotes. Macro variables can be especially problematic as they will generally not resolve when they are inside single quotes. However this is SAS, and that implies that there are several things going on at once and that there are several ways to solve these types of quoting problems. The primary goal of this paper is to assist the programmer with solutions to the quotes within quotes problem with special emphasis on the presence of macro variables. The various techniques are contrasted as are the likely situations that call for these types of solutions. A secondary goal is to help the reader understand how SAS works with quote marks and how it handles quoted strings. Although we will not go into the gory details, a surface understanding can be useful in a number of situations.

Paper CC04

Monday: 9:30 - 9:40 a.m.

Room: Sapphire D

---

### **Standardization of Confidence Intervals in PFS Tables - a Macro Approach**

Author(s): John Saida Shaik, Boxun Zhang

Any oncology trial requires time-to-event (TTE) analysis to determine whether an event of interest (EOI) occurred and when it occurred. An EOI can be progression of disease, stable disease, complete remission or death. Usually, overall survival (OS) and/or progression-free survival (PFS) tables are requested by statisticians to perform TTE analyses. In this paper, we propose a macro that uses the Proc Lifetest procedure to automate generation of PFS tables and standardize calculation of confidence intervals (CI), in addition to a plethora of other parameters. Usage of Proc Lifetest to generate PFS tables results in consistency. This macro makes calculation of CI much more straight-forward and a standardized process.

---

Paper CC05

Monday: 9:45 - 9:55 a.m.

Room: Sapphire D

### **Having Fun with RACE Derivation in DM Domain**

Author(s): Chunxia Lin, Deli Wang

Race in SDTM DM domain is an expected variable; generally it is quite easy to derive. However, there are times when raw database is set up with multiple race variables (RACE\_X), and a subject may have multiple race values. Some clients require that, if a subject has more than one race values, set RACE as "MULTIPLE", otherwise, set to the single selected race value. The logic looks simple, however SAS programmers have to figure out two important steps before coding: 1. whether the subject has more than one race variables selected; 2. which race variable has nonmissing value. The Multiple IF-THEN/ELSE statements work fine but the coding is a little tedious. This paper introduces four methods to derive RACE using the logical expression and SAS functions. Each of the methods is discussed separately depending on the attributes of the data (character/numeric).

---

Paper CC06

Monday: 10:15 - 10:25 a.m.

Room: Sapphire D

### **Tips to Manipulate the Partial Dates**

Author(s): Deli Wang, Chunxia Lin

A partial date is simply any date where the date is incomplete, but not wholly missing. More commonly in clinical trials, the day and/or month are missing. In these cases, SAS programmers may be asked to impute a reasonable date or time per client's requirement or statistical purpose. This paper introduces two different imputation logics with the missing day set to the last day of the month. Both methods take the leap years into consideration, and generate same results. However, one method imputes the day depending on the leap year and month, while another not considering any.

---

Paper CC07

Monday: 10:30 - 10:40 a.m.

Room: Sapphire D

### **Preserving Line Breaks When Exporting to Excel**

Author(s): Nelson Lee

Do you have imported data with line breaks and want to export the data to Excel? What if you need to preserve the line breaks in the data export? This paper shares a quick tip on how to export SAS data with line breaks to Excel through PROC TEMPLATE and TAGSET attributes. This paper was written for audiences with beginner skills; the code was written using SAS® version 9.2 on Windows operating system.

Paper CC08

Monday: 1:30 - 1:40 p.m.

Room: Sapphire D

**Additional Metadata for Common Catalog Entry Types**

Author(s): Ken Borowiak

Anyone who has worked with SAS® has probably added descriptive attributes to entities such as variables (labels and formats), data sets (labels), reports (titles and footnotes), and programs (comments). Though not as well known, one can add descriptive labels to commonly used catalog entry types, namely formats and macros. This paper will demonstrate how to add, modify and retrieve this metadata.

---

Paper CC09

Monday: 10:45 - 10:55 a.m.

Room: Sapphire D

**An alternative way to detect invalid reference records in supplemental domains**

Author(s): Steven Wang

Quite often in the process of development of SDTM datasets, we may just be interested in how to identify some specific errors in certain domains without the time and effort of running the OpenCDISC validator.. The purpose of this paper is to demonstrate an alternative way to identify the invalid cross-referenced records with error code SD0077 in the supplemental domains by utilizing the SAS macro language.

---

Paper CC10

Monday: 11:00 - 11:10 a.m.

Room: Sapphire D

**Automating the Number of Treatment Columns for a Dose Escalation Study**

Author(s): Sonali Garg, Catherine Deverter

In dose escalation studies, there is the need to increase or decrease the number of treatment columns in table output when a new treatment arm is added to the trial. If programs are not written to add the new columns dynamically, they often need to be edited repeatedly. What if we can automate the number of columns to be displayed in the table without having to go and edit the program? This paper will show how SAS programmers can automate and control the number of columns in the table output with having to change the PROC REPORT's COLUMN and DEFINE statements again and again. The method discussed in this paper will make use of the SQL procedure.

---

Paper CC11

Monday: 11:15 - 11:25 a.m.

Room: Sapphire D

**Streamline the Dual Antiplatelet Therapy Record Processing in SAS by Using Concepts of Queue and Run-Length Encoding**

Author(s): Kai Koo

Dual antiplatelet therapy (DAPT) is a common practice to protect patients from stent thrombosis after stent implantation. In order to analyze its efficiency, medication records of two antiplatelet drugs have to be compiled together. To reduce the coding complexity, the concept of queue is used in SAS data steps. In this approach, multiple records of medication started and stopped status for single patient are rearranged as a "first in, first out (FIFO)" structure through building a new dataset dynamically. Under this queue programming logic, the complete medication record can be derived easily and also compacted into a readable medication history variable simultaneously by using the approach of "run-length data compression".

---

Paper CC12

Monday: 11:30 - 11:40 a.m.

Room: Sapphire D

### **A Macro to Automate Symbol Statements in Line Plots**

Author(s): Deli Wang, Chunxia Lin

When processing PROC GPLOT procedure to generate line plots, SAS programmers may have to write several symbol statements and customize specific colors for different groups. The process might be very annoying when data have multiple groups, or programmers are asked to generate/update tons of plots within tight timelines. So isn't wonderful if the symbols in line plots can be assigned based on the input data and line attributes defined in data step, hence programmers take accurate control of line attributes? This paper is trying to write a small macro to empower the programmers with an easy control of symbol statements. The macro provides several advantages: 1. Accurately assign the line attributes like writing a datastep; 2. No longer need to figure out how many symbol statements needed and which symbol is for which line; 3. Can easily modify the macro to fit other plotting needs, program once and use it always.

---

Paper CC13

Monday: 11:45 - 11:55 a.m.

Room: Sapphire D

### **WHERE, Oh, WHERE Art Thou? A Cautionary Tale for Using WHERE Statements and WHERE= Options**

Author(s): Britney Gilbert

Using WHERE statements in your SAS programming can make processing data more efficient. However, the programmer needs to be aware of how SAS processes multiple WHERE statements or combinations of WHERE statements and WHERE= options in data steps and procedures. This paper explores examples of the uses of WHERE statements and WHERE = options in both data steps and procedures and the resulting logs and output.

---

Paper CC14

Monday: 3:00 - 3:10 p.m.

Room: Sapphire D

### **A Toolkit to Check Dictionary Terms in SDTM**

Author(s): Huei-Ling Chen, Helen Wang

WHO Drug Dictionary (WHODD) and MedDRA are two important dictionaries used in the SDTM datasets. In the intervention domains, WHODD is used to derive the standardized medication name. In the event domains, the standardized text description of an event is based on dictionary MedDRA. A standardized dictionary term can be derived from an original term when its value exists in the dictionaries. Sometimes an original term fails to facilitate coding due to its value does not exist in the dictionary. In a busy project team, preparing SDTM packages for multiple studies is very common. It will be helpful if there is a checking toolkit to quickly identify records that fail to have dictionary term derived when preparing SDTM packages. This paper will first briefly summarize the application of WHODD and MedDRA and list out the dictionary-derived variables commonly used in clinical trial studies. Then a method to check every SDTM domains with dictionary WHODD or MedDRA derived variables is described here. The method is written into a macro format which can check the entire SDTM package without pre-requisite knowledge on which domain has dictionary derived terms. Another benefit of this handy toolkit is that it is a portable macro that can easily be adapted and adopted in other applications.

Paper CC15

Monday: 2:30 - 2:40 p.m.

Room: Sapphire D

**Cleaning up your SAS® log: Overwritten Variable Info Messages**

Author(s): Jennifer Srivastava

As a SAS programmer, you probably spend some of your time reading and possibly creating specifications. Your job also includes writing and testing SAS code to produce the final product, whether it is SDTM datasets, ADaM datasets or statistical outputs such as tables, listings or figures. You reach the point where you have completed the initial programming, removed all obvious errors and warnings from your SAS log and checked your outputs for accuracy. You are almost done with your programming task, but one important step remains. It is considered best practice to check your SAS log for any questionable messages generated by the SAS system. In addition to messages that begin with the words WARNING or ERROR, there are also messages that begin with the words NOTE or INFO. This paper will focus on the overwritten variable INFO message that commonly appears in the SAS log, and will present different scenarios associated with this message and ways to remove the message from your log, if necessary.

---

Paper CC16

Monday: 1:45 - 1:55 p.m.

Room: Sapphire D

**Need for Speed in Large Datasets - The Trio of SAS® INDICES, PROC SQL and WHERE CLAUSE is the Answer**

Author(s): Kunal Agnihotri

Programming on/with large datasets can often become a time taking ordeal. One way to handle this type of situation is by using the powerful SAS® INDICES in conjunction with the WHERE clause in the PROC SQL step. This paper highlights how effective indices can be created using SQL (more flexible when compared to indices created using DATA step index option or the DATASETS procedure) and further subset data using the WHERE clause which drastically reduces dataset processing and run time. This combination of techniques gives better results when accessing big datasets than using the said techniques singularly. This paper also throws light on the dual functionality of the SQL created indices viz., the ability to create indices on both new and existing SAS datasets.

---

Paper CC18

Tuesday: 11:45 - 11:55 a.m.

Room: Sapphire D

**Macro to check Audit compliance and standards of SAS programs**

Author(s): Seeja Shetty

As a lead statistical programmer in clinical trial study, you are responsible to submit an accurate representation of the collected data in the form of analysis datasets, summary tables, listing and figures. In addition, you need to ensure the SAS programs which created the above mentioned reports are compliant to Good Programming Practices. Even considering a small study it is tedious and time consuming to go through all the programs individually. This paper demonstrates a SAS macro to ensure basic checks for the GPP compliance which might divulge inconsistencies during study Audits. Products used: SAS 9.2 and MS Excel 2007  
Operating System: Windows XP Skill level: Intermediate

---

Paper CC19

Monday: 3:30 - 3:40 p.m.

Room: Sapphire D

**Let Chi-Square Pass Decision to Fisher's Programmatically**

Author(s): Linga Reddy Baddam, Sudarshan Reddy Shabadu

There is often the question of "when" and "which" statistical test should be used against any kind of clinical data in wide variety of situations, as there are many significant cases when insufficient information can lead a statistician to choose an inappropriate analysis, it is common to test the hypothesis of independence of events across different treatment arms, and identify the appropriate statistical test based on certain clinical reporting criterion. The clinical programmer has to develop code to carry out the subsequent statistical tests based on the nature of data and assumptions which have been made. To minimize and optimize efforts of clinical programmers, automating code to find the appropriate statistical test is key-job, the features of the %fc\_pval macro explicitly discussed further in detail.

---

Paper CC20

Monday: 3:45 - 3:55 p.m.

Room: Sapphire D

**BreakOnWord: A Macro for Partitioning Long Text Strings at Natural Breaks**

Author(s): Richard Addy, Charity Quick

Breaking long text strings into smaller strings can be tricky, if splitting the string needs to be based on something more than simply length. For example, in SDTM domains, character variables are limited to 200 characters. Excess text need to be placed into supplemental datasets (and the variables there are also limited to 200 characters) - but it's not sufficient to just break the text into 200-character chunks; the text needs to be split between words. This paper presents the BreakOnWord macro, which breaks a long text variable into a set of smaller variables of a length specified by the user. The original text is partitioned text at natural breaks - spaces, as well as user-supplied character values. The macro will create as many variables as necessary, and the variables are named using a user-supplied prefix and an ascending suffix. The user has the option of naming the series of variables beginning with the prefix only (creating SDTM-friendly variable names: AETERM, AETERM1, AETERM2, for example). BreakOnWord checks that the user inputs are valid (the specified data set exists, and the specified long text string is present in that data set) and that the variables to be created by the macro are not already present. The newly created variables are added to the input data set, or, optionally, the macro can output a new data set. BreakOnWord requires SAS of 9.1 or higher.

---

Paper CC21

Monday: 4:00 - 4:10 p.m.

Room: Sapphire D

**Automating Production of the blankcrf.pdf**

Author(s): Walter Hufford

The blank Case Report Form (blankcrf.pdf) is a critical component of the NDA submission. Per FDA guidance, source data domain, variable name and controlled terminology for each case report form (CRF) item included in the tabulation datasets submitted should be displayed on the blankcrf.pdf. Production of the blankcrf.pdf is a tedious, non-programming task that is increasingly becoming the responsibility of the statistical programmer. This paper describes an easy to use, automated method of annotating the CRF.

---

Paper CC22

Monday: 4:15 - 4:25 p.m.

Room: Sapphire D

### **A SAS Macro Tool for Visualizing Data Comparison Results in an Intuitive Way**

Author(s): Hui Wang, Weizhen Ying

In clinical data analysis, PROC COMPARE is widely used in data quality control. However, sometimes the results of this procedure can be quite challenging to understand. For example, it only displays the first 20 characters in value comparison output. Consequently, character mismatches in lengthy text strings may not be seen directly. In addition, it shows mismatches separately and does not simultaneously show values of variables that might be relevant to mismatches, which leads to difficulty in figuring out why mismatches happen. This paper intends to introduce a SAS macro tool which is based on PROC COMPARE but gives an intuitive comparison report. First of all, this macro juxtaposes mismatched values vertically and thus allows the display of whole variable content. Secondly, the macro can extract unique rows from datasets compared and then tells directly which rows are missing in which dataset. Finally, it enables the display of values of pre-specified variables along the same row of mismatches. In this way, the programmer can intuitively observe the possible involvement of other variables in causing mismatches. In short, this macro, with its user-friendly output, functions as an extension tool of PROC COMPARE and is able to improve efficiency of clinical data analysis.

---

Paper CC23

Monday: 4:30 - 4:40 p.m.

Room: Sapphire D

### **Need to Review or Deliver Outputs on a Rolling Basis? Just Apply the Filter!**

Author(s): Tom Santopoli

Wouldn't it be nice if all of the outputs in a deliverable passed QC at exactly the same time and could be submitted for final review all at once, while leaving adequate time for the review to be completed prior to a deadline? Unfortunately, outputs often pass QC at different times and must be submitted for final review on a rolling basis in order to meet deadlines. The task of selecting 50 specific outputs to copy from a folder containing 500 outputs can be very tedious and cumbersome for a lead programmer. There are papers that explain how to copy files from one folder to another, but this paper addresses the issue of selectively copying specific files based on criteria set in a project tracking document. A simple macro called %MFILTER is presented to help make life a little easier for lead programmers as deadlines approach.

---

Paper CC25

Monday: 2:00 - 2:10 p.m.

Room: Sapphire D

### **SCAN and FIND "CALL SCAN"**

Author(s): Usha Kumar

Let's explore the unexplored. Most of us are frequent users of SCAN and FIND/INDEX function. We use these when it comes to parsing a string. If we were to figure out the nth word from a character string, we would do so by using SCAN to solve this. If we are to figure out the starting position of the nth word extracted, we would try and use INDEX/FIND for this purpose. What if we come up with a situation where we are to find out the position of the nth word that is repeated multiple times in a string? It gets more complicated. We have definitely come across this situation and solved it too. But let's find out if it is as easy as solving using the CALL SCAN routine available in SAS v9.1 and above.

Paper CC29

Tuesday: 8:00 - 8:10 a.m.

Room: Sapphire D

### **Quickly Organize Statistical Output for Review**

Author(s): Brad Danner

Statisticians and programmers in the pharmaceutical industry are often required to include model based tables to satisfy the requirements of study protocols. Working for a client who does not explicitly require or expect the output created from SAS procedures as part of a submission has understandably created difficulty when reviewing such tables for accuracy. The programming team tends to focus on the end product with less thought as to checking the SAS procedural output. However, statisticians reviewing and approving the table will often require the statistical output to determine if the procedure was applied correctly. If prompted to produce the statistical output, a flood of information is typically consigned to a large text file, often with little or no organization which then impedes review. A recommended alternative employing judicious use of titles and options is proposed here to improve the readability and review of the statistical output for both programmers and statisticians.

---

Paper CC30

Tuesday: 8:15 - 8:25 a.m.

Room: Sapphire D

### **Times can be Tough: Taming DATE, TIME and DATETIME variables**

Author(s): Sajeet Pavate

Some programmers may not fully understand how the values for Date, Time and Datetime variables are stored and manipulated within SAS especially in relation with each other. Several previous papers have provided an introduction to how date, time and datetime values work in SAS as well as the different functions and formats that apply to these variables. The aim of this paper is to show common issues that may occur while working with these variables. These issues, for example, occur when missing time component is not considered or when incorrect assumptions are made that can result in invalid or inaccurate calculations. The paper points to some of these issues with examples from real life scenarios and includes the corrected code to fix such issues. It attempts to educate readers on how these variables are used within SAS and makes them aware and mindful of common pitfalls when working with Date, Time and Datetime variables.

---

Paper CC31

Tuesday: 8:30 - 8:40 a.m.

Room: Sapphire D

### **Let the CAT Catch a STYLE**

Author(s): Yanhong Liu, Justin Bates

Being flexible and highlighting important details in your output is critical. The use of ODS ESCAPECHAR allows the SAS® programmer to insert inline formatting functions into variable values through the data step, and it makes a quick and easy way to highlight specific data values or modify the style of the table cells in your output. What is an easier and more efficient way to concatenate those inline formatting functions to the variable values? This paper shows how the CAT functions can simplify this task.

---

Paper CC32

Tuesday: 8:45 - 8:55 a.m.

Room: Sapphire D

### **1 of N Methods to Automate Y-axis**

Author(s): Sanjiv Ramalingam

Methodologies for automating the Y-axis have been discussed by many previously. The motivation to share this code is to not only present readers an another algorithm but because the methodology is unique, simple, robust in that it can be used not only with integer data but also with floating point data and has comparatively very few lines of code. The algorithm also allows the user to specify in how many divisions they would like their data to be visualized. Even though SAS 9.3 handles Y-axis automation, reasons as to why automation by the user is warranted in certain situations have been elucidated.

Paper CC33

Tuesday: 9:00 - 9:10 a.m.

Room: Sapphire D

### **QC made Easy using Macros**

Author(s): Prashanthi Selvakumar

To err is human. But this cannot be an excuse for our mistakes. We have to make sure we have no errors in our reports that we submit. To minimize errors in programming, we have the QC (Quality Check) team. In most cases, QC requires producing reports from the scratch, which is time consuming. In this paper we will be discussing a macro used for creating reports similar to the production programmer and that compares the result with QC programmer. This can help you save time on coding as well check the mistakes that human eyes do not capture.

---

Paper CC36

Tuesday: 9:30 - 9:40 a.m.

Room: Sapphire D

### **Macros make Final Documentation Quick and Easy**

Author(s): Indrani Sarkar, Jean Crain

Purpose of this macro is to generate a document containing information on all the sas programs used to create derived datasets, summary, listings, and figures during a clinical trial project. This word document is used as a part of final package delivered to the client after completion of each project. Primary objective for this file is to assemble all the key information in one document so that it would be easier to replicate all the work in future without any difficulties and also it will help someone to get familiar with a project in a very short time. The macro, explained in this paper, will be able to list the sas program names; the analysis population(s); the endpoints; name(s) of required input datasets (including non-sas files); any additional SAS macros called by the program; the output file name(s); and any additional notes clarifying information. It would also be able to group tables, listings and figures based on their endpoints. This macro uses PRVF (Programming Review and Validation form), an excel workbook which keeps track of deliveries made at different points of time, and all programs as source files. Our expectation is that this macro will reduce number of hours usually spent on the documentation as well as human errors and would bring consistency in format across different projects.

---

Paper CC37

Tuesday: 9:45 - 9:55 a.m.

Room: Sapphire D

### **Reducing Variable Lengths for Submission Dataset Size Reduction**

Author(s): Sandra Vanpelt Nguyen

The FDA has cited dataset size as one of the issues they commonly encounter for submitted clinical trial datasets and has found the allotted column variable lengths to have a high correlation to overall dataset size. Based on this analysis, the reviewing divisions have requested sponsors to reduce variable lengths to the minimum lengths needed to accommodate the values found within each variable. Since it may be difficult to identify or predict up front the longest potential value for every variable, not to mention that sponsors will want to avoid risk of truncation, this paper presents a macro which identifies the minimal lengths needed based on the actual data values which are present in a dataset and reassigns variable lengths accordingly as a post-processing step prior to submitting datasets to a regulatory agency.

---

Paper CC38

Tuesday: 10:15 - 10:25 a.m.

Room: Sapphire D

### **Let SAS Do That For You**

Author(s): Emmy Pahmer

Do you ever repeat certain tasks and think there should be a better or easier way to do them? Choosing and copying files are tasks that we often perform manually but which can easily be done by SAS using particular selection criteria. We'll look at how to get a list of files in a directory, some code to select the desired file, creating a warning message if more than one file fits the criteria or if no files fit the criteria, and copying the file(s). This presentation is suitable for all users, including beginners.

---

**Inserting MS Word Document into RTF Output and Creating Customized Table of Contents Using SAS and VBA Macro**

Author(s): Haining Li, Hong Yu

In clinical trials, to effectively monitor study progress and subject safety, the reviewers frequently request to submit a set of tables and listings at regular basis throughout the study. With the ODS (Output Delivery System) RTF destination, the programmers build tables and listings that are opened directly in MS Word and other word-processing packages. To expedite the review process, with each submission, a Summary Report to highlight any study update during the report cycle is highly recommended. In practice, the programmers are often provided with a MS Word document with multiple sections to be inserted into the RTF outputs. Having all information consolidated in one document will keep the integrity of the report and make it easy for reference. Also, with the various length of the Summary Report at each submission, it would be helpful if the SAS program could automatically create a Table of Contents (TOC) which includes the hyperlinks to each section of the Summary Report in the final RTF outputs. Unfortunately, SAS does not provide a stable function to insert the MS Word document into the RTF outputs and create the customized TOC. This paper will offer a method to insert any MS Word document and prepare the inserted file to create customized TOC using SAS and VBA Macro. With this approach, the customized TOC will maintain the hyperlinks to configurable sections of the inserted documents. This method is very flexible, robust and can find broad application in SAS reporting.

**PRELOADFMT comes to your rescue, it brings missing categories to life in summary reports**

Author(s): Niraj Pandya, Ramalaxmareddy Kandimalla

In clinical trials data analysis and reporting, one always comes across producing categorical tables used for analysis. Provided with lots of different kind of data types such as labs, adverse events, physical examination, concomitant therapies etc and multiple categories for several categorical variables, it becomes difficult sometimes to impute 0 values for categories which are missing altogether for a particular variable. SAS provides multiple options to handle such cases and produce output with all the categories even if some don't exist in the data. This paper concentrates on different work around methods and will discuss details of implementing each technique.

**Give me everything! A macro to combine the CONTENTS procedure output and formats.**

Author(s): Lynn Mullins

The PROC CONTENTS output displays SAS® data set information such as variable name, type, length, informat, and format names. The values and codes for the formats are not included in this output; therefore, a separate printout of the program that creates the formats or a printout of the format catalog via data step programming needs to be used in conjunction with the PROC CONTENTS output. This paper will describe a SAS® macro that combines PROC CONTENTS output with the format catalog data to create a single metadata dictionary file.

Paper CC44

Tuesday: 11:15 - 11:25 a.m.

Room: Sapphire D

### **A Macro to Create Occurrence Flags for Analysis Datasets**

Author(s): Ed Lombardi

The different paths available for occurrence analysis are fairly straightforward. The path that takes a little longer ends up giving a better end product. This approach is to create occurrence flags in analysis datasets as described in ADaM's ADAE structure. These flags allow for standardized table programs while also providing traceability to allow reviewers to know what records are being used in analysis. The code to create these flags is straightforward and repetitive so it's an excellent situation to use a macro. Otherwise, creating these flags can lead to cluttered sections of analysis dataset programs.

---

Paper CC45

Tuesday: 11:30 - 11:40 a.m.

Room: Sapphire D

### **Self-fulfilling Macros Generating Macro Calls and Enabling Complete Automation**

Author(s): Y. Christina Song

In most cases, macros facilitate repetitive SAS programming iterations. In clinical data cleaning tasks, like edit check programming, each query item has different query logic. However, the overall query report generation process is somewhat repetitive; it all includes the process of reading, subsetting, formatting, and printing data. Even with all of the macros, programmers still have to read specifications carefully and create macro calls accordingly. The overall process is time-consuming and labor intensive. To improve programming efficiency, a procedure is introduced to let SAS macros generate macro calls from the specification sheet and to do all of the SAS programming. This program takes advantage of the data-driven and dynamic features of SAS macros and dynamically reads specifications, tweaks the data, generates all macro calls, codes the texts between the macros, formats the data, and outputs the data into desired reports. This paper outlines the key elements and basic steps of the macros, and discusses how this strategy could be used to create other macros generating macro calls and enabling automatic operations. It may also be used for similar tasks that come with a specification sheet, such as generating some of standard analysis data sets.

---

Paper CC47

Monday: 2:15 - 2:25 p.m.

Room: Sapphire D

### **It's not the Yellow Brick Road but the SAS PC FILES SERVER® will take you Down the LIBNAME PATH= to Using the 64-Bit Excel Workbooks.**

Author(s): William E Benjamin Jr

SAS users who also use Excel or produce Excel workbooks will eventually find that the rapid pace of hardware and software changes occurring today will soon meet them head on. The need for faster computers and bigger workbooks are accelerating. This issue is not new, 8-Bit computers were replaced with 16-Bit CPU's, and they lost the battle to 32-Bit computers. Today new computer hardware routinely comes with up to eight 64-Bit CPUs in a single 2X2 piece of hardware for an affordable price. The fictional path down the yellow brick road to OZ was fraught with many challenges. SAS Institute has stepped up to the challenge of the way their users mix and match hardware and software. Since SAS Institute could not control the users and their hardware, the institute expanded the way programmers can use the LIBNAME statement to create a PATH= to open the way to create and read new Excel formats. Some new interfaces were built to facilitate transfers, one interface called the PCFILES server allows passing data to and from SAS and Excel across the boundary between 32 and 64-bit computers and software. The examples presented here will clear the fog and open the doorway past the curtain to view the processes available to open a PATH= to become your company's SAS to Excel wizard.

---

**Creating PDF Reports using Output Delivery System**

Author(s): Shubha Manjunath

Output Delivery System (ODS) in V9.2 provides new and enhanced capabilities to programmers for reporting and displaying clinical trials results with numerous options that provide greater control of formatting and layout of the report. The purpose of this paper is to list and demonstrate variety of ODS parameters that help control display of statistical reports while creating PDF reports. Fundamental ODS procedures included in the article but not limited to are, controlling the structure of the report (eg. ODS DOCUMENT), sending reports directly to a printer (eg. ODS PDF), inserting text into an ODS output (ODS TEXT =), setting valid values for page orientation (eg. PAGESYTL) and controlling the level of expansion of PDF table of Contents (eg. PDFTOC). These functionalities are illustrated using SAS9.2 version on a Windows operating system. The objective of this paper is to reveal techniques that were used to produce ODS reports from 'Legacy' reports - 'Legacy' throughout this presentation refers to non-ODS practice (practice that was in use before the implementation of ODS reporting). This paper does not provide a preface to Proc Report or ODS, but provides some easy and realistic guidelines to produce ODS reports and help minimize time and effort on the part of Statistical programmers and Statistical users. It provides some realistic guidelines with an emphasis on illustrations of basic ODS functionalities that would require minimal syntax usage. These options will be illustrated with simplified examples and case studies.

**A Shout-out to Specification Review: Techniques for an efficient review of Programming Specifications**

Author(s): Sajeet Pavate, Jhelum Naik

One of the most common reasons for poor quality of programming output, higher costs due to large number of programming hours and missed timelines is poorly written programming specifications (specs). A thorough independent review of the specifications by competent programmers and/or statisticians prior to programming activities is an important safety net to ensure the specs are in a suitable form for programming. This early intervention eliminates the risks to deliverable in terms of quality, costs and timelines. This paper provides clear guidelines, tips and techniques to empower spec reviewers to perform an efficient spec review. This paper provides techniques for spec review that can be implemented for different types of deliverables such as Transformed/Mapped Data, Analysis Database, Tables Listings and Figures and for different types of organizations such as CROs and pharmaceutical companies. The techniques discussed in this paper can also be adapted and used in organizations which have built-in standard specs - it empowers the foot soldiers on the project team to identify and raise concerns that may not be evident in standards generated by a working group or standards committee.

**What do you mean 0.3 doesn't equal 0.3? Numeric Representation and Precision in SAS and Why it Matters**

Author(s): Paul Stutzman

The particular manner in which numeric values are stored can cause SAS programs to produce surprising results if it is not well understood. Numeric representation and precision describe how these values are stored. This paper will show how numeric representation and precision can affect program output and produce unintended results. It shows how numbers are actually represented, it identifies the magnitude of the difference between how values are represented and their absolute values, and it provides solutions to the problems that numeric representation and precision can cause.

## Career Planning

---

Paper CP01

Monday: 3:30 - 4:20 p.m.

Room: Exhibit Area - Demo Theater

### **What's Hot, What's Not - Skills for SAS® Professionals**

Author(s): Kirk Paul Lafler, Charlie Shipp

As a new generation of SAS® user emerges, current and prior generations of users have an extensive array of procedures, programming tools, approaches and techniques to choose from. This presentation identifies and explores the areas that are hot and not-so-hot in the world of the professional SAS user. Topics include Enterprise Guide, PROC SQL, JMP, PROC REPORT, Macro Language, ODS, DATA step programming techniques such as arrays and hash, sasCommunity.org®, LexJansen.com, JMP®, and Output Delivery System (ODS).

---

Paper CP02

Tuesday: 4:00 - 4:20 p.m.

Room: Exhibit Area - Demo Theater

### **Show Your SAS® Off: Personal Branding and Your Online Portfolio**

Author(s): Barbara Ross

The word "online portfolio", for most people, brings to mind collections of stunning photographs and graphic design; though utilization of this tool should not stop there. Any and all creative professionals can and should have an online portfolio. The ferocity at which we present and attend these user groups attests to the fact that portfolios would flourish within our line of work. Are white papers not but showcases of our creative programming solutions? How many of us keep copies of our best coding snippets tucked away in a file on the computer? Would it not serve us best to showcase these to our peers? After all visibility creates opportunities. Opportunities that include new job leads, collaborations, mentorship, and self-learning. Also it's simple. In today's day and age, sites such as Weebly and Wordpress can have you published within an hour- all you need is content and a little inspiration. This paper hopes to provide that inspiration by giving examples of successful online programming portfolios and tips on how to create your own.

---

Paper CP03

Monday: 4:30 - 5:20 p.m.

Room: Exhibit Area - Demo Theater

### **Creating a personal career GPS in a changing workplace: Ten steps to improve your professional value.**

Author(s): Bill Donovan, Shridhar. Patel

Building a successful career in the life sciences field requires that you regularly measure and strengthen your professional value. But personal development is particularly challenging within this sector. The industry thrives on the promise of breakthrough drugs, but the lack of clarity about formulations and the variable cost structure of clinical trials means many seasoned professionals perform in outsourced or contract roles. In this lean, results-focused corporate culture, companies are unwilling or unable to offer career planning resources to the professionals who contribute to pharmaceutical and biotechnology success. Yet the employment boom continues. Certain industry estimates project 30 percent growth through 2016, so hiring remains robust. Within this specialized workforce, a paradigm is shifting. Contract work has become the leading form of employment, the new normal. Now more than ever, you need a detailed plan designed to help drive your career forward. We outline these steps within Creating a personal career GPS in a changing workplace: Ten steps to improve your professional value progressively by strengthening your career strategy. In this paper and presentation, we challenge life sciences professionals to take charge of their personal development and help them navigate toward sustained success. It will not happen accidentally! Life sciences talent must respond to this trend by crafting a customized career strategy. Individuals need a guide, a sort of personal career global positioning system (GPS), to identify opportunity and structure their career progression effectively.

---

Paper CP04

Tuesday: 3:30 - 3:50 p.m.

Room: Exhibit Area - Demo Theater

### **Pharma Company Questions and Answers**

Author(s): J.J. Hantsch

We've all heard that pharmaceutical companies have two factions, a science directorate that doesn't care about sales and a sales directorate that doesn't understand the science, but there is a third group, the Slartibartfast team. But who are they? Are new hire statisticians any better equipped to deal clinical trials than new hire statistical programmers? Why is the CRA function the only one to avoid when considering dating within your company? What is the difference between project teams and product teams? How did CROs come about? Does anyone ever really adopt an orphan drug? What is so different about the medical writers? What is the difference between PK and PD? Why are the largest pharmaceutical companies in the world mostly US companies? Will MedDRA ever run out of numbers and revert to COSTAR? These and other mysteries of your average pharmaceutical company answered today. Answers to these and other mysteries, along with un-common tidbits of your average pharmaceutical company will be discussed.

---

Paper CP05

Monday: 1:15 - 1:35 p.m.

Room: Exhibit Area - Demo Theater

### **Call me, Maybe? Using LinkedIn to make sure you get the call for the job**

Author(s): Tara Potter, Adel Lesniak

LinkedIn is one of the world's most popular social networking websites and is the leading networking site among professionals. Whether you are job searching or not, it is critical that you understand the importance of a well written profile and how exactly recruiters and other hiring professionals use this information when searching for top talent. It goes well beyond the basics of your job title, employment history and whether or not you have applied to a posted position. LinkedIn is an excellent tool to help you position yourself accurately and one that recruiters are now using more than ever to find top talent for their organizations!

---

Paper CP06

Tuesday: 1:15 - 2:05 p.m.

Room: Exhibit Area - Demo Theater

### **Enhancing Your Career by Bringing Consultants into Your Organization**

Author(s): Roger Muller, Gregory Nelson

No matter how outstanding you are as programmer, your value to the organization that employs you is based upon your perceived contributions to the overall goals and success of the organization. This paper addresses the role that consultants (both external and internal) can bring to an organization, thereby enhancing not only your performance, but that of others. Issues are presented for both the client and consultant. Consultants can have a larger view than just that of the task at hand. They can play a vital role in helping establish the direction to be taken on not only current work, but also on future projects. In the process, the client associated with bringing the consultant into the organization will be perceived as a strong team player. It is essential that the consultant and the sponsor be proactive in broadcasting their accomplishments and make higher levels of management aware of progress. Several specific examples will be presented where the presenters were in one of these roles. The examples relate to IT projects, not all of which were SAS. In each example both the consultant and the sponsor bringing them in was perceived well.

---

Paper CP07

Tuesday: 2:15 - 3:05 p.m.

Room: Exhibit Area - Demo Theater

**Careers in Biostatistics and Clinical SAS® Programming - An Overview for the Uninitiated**

Author(s): Justina Flavin

In the biopharmaceutical industry, biostatistics plays an important and essential role in the research and development of drugs, diagnostics, and medical devices. Familiarity with biostatistics combined with knowledge of SAS® can lead to a challenging and rewarding career that also positively impacts and transforms patients' lives. This presentation will provide a broad overview of the different types of jobs and career paths available, discuss the education and skill sets needed for each, and present some ideas for overcoming entry barriers into careers in biostatistics and clinical SAS® programming.

---

Paper CP08

Tuesday: 4:30 - 4:50 p.m.

Room: Exhibit Area - Demo Theater

**Pick Me, Pick Me**

Author(s): Deloris Jones

Hiring managers and programmers alike are experiencing the same conundrum. Managers want the highest quality programmers for the lowest price possible and programmers want the highest pay attainable with the most prestigious employer. We are currently experiencing high demand and low supply in a volatile market where CEO's are Outsourcing, Offshoring, or keeping the work In-house based on cost saving initiatives and forecasting of their pipeline. The hiring approach is slightly different for Sponsors and CRO's however, because these managers typically want and need a flexible headcount model, which entails engaging programmers on an ad-hoc basis. Essentially, it all boils down to quality and pricing. The industry trend is to get and do more with less. Some glaring challenges we repeatedly face include: Pricing Disparities, Client Rigidity and Candidate Flexibility. " Pricing Disparities are evident when you have two equally desirable biotech companies located within 50 miles of each other and are seeking a contract statistical programmer with 5 years industry experience but have a 20% variance in contractor bill rates. " Client Rigidity varies widely depending upon the client. For instance, a client needs a contractor who is proficient in R and S+ and now we find ourselves searching for the "purple squirrel" (a very rare candidate). Later we find out that R or S+ experience is fine. " Candidate Flexibility is multifaceted including but not limited to rate, location and/or telecommuting.

---

Paper CP09

Monday: 1:45 - 2:05 p.m.

Room: Exhibit Area - Demo Theater

**Career planning-How to make it work for you? - Tips for programmers, managers and senior leaders**

Author(s): Vijay Moolaveesala

Most employees depend on their employer's processes to define their career progress and career ladder. Realizing these expectations, many organizations have taken the responsibility of investing in employees' career planning. But, do we need to depend on others for our career growth? Do we need to depend on management or organization's initiative to plan our career? Do we feel that these initiatives are producing the results to the satisfaction of associates? Do we use them effectively? This paper will provide tips to approach career planning which can result in individual employee satisfaction and organization wide positive impact. This paper describes 3 step process of career planning (Vision Statement, Goal Setting and Action Plan) and ways to review, seek feedback and manage career progression. This paper also provides tips to manage career plans for high performing individuals along with tips for individuals to manage their own career. Career planning is not just for associates who aim for meteoric rise but for all who aspire to advance in their career. If career planning starts with a notion of title advancements, then it is destined to disappoint many associates. Career planning is an ongoing process of evaluating and reviewing career objectives and adjusting the approach to align with personal & professional priorities as well as with job market changes. Most of the employees may be having career progression discussions with their supervisors but as long as these are not approached with documented process, over the time this initiative will lose its steam.

---

**Negotiation: Getting the Best Out of an Offer**

Author(s): Janet Stuelpner

The job offer has come in, maybe even more than one. You need to know the types of things that are worth discussing and on which terms you should negotiate and those that are not worth mentioning. There are many things that are contained in an offer, but there are sometimes things that are not mentioned that may be important to you. This paper will show the pros and cons of an offer and what is significant and worth discussing with your potential employer.

---

## Data Visualization and Graphics

---

Paper DG01

Monday: 3:30 - 3:50 p.m.

Room: Sapphire L

### **Techniques of Preparing Datasets for Visualizing Clinical Adverse Events**

Author(s): Amos Shu

With the introduction of SAS® Graph Template Language and SAS/GRAPH® SG Procedures, there have been papers [1-4] that have provided very useful codes to easily visualize clinical adverse events (AE). However, it is still not easy to create the suitable datasets that are used to generate the right graphs. This paper discusses techniques of preparing datasets that will be used to generate different types of figures for clinical adverse event.

---

Paper DG02

Monday: 4:00 - 4:20 p.m.

Room: Sapphire L

### **Graphical Representation of Patient Profile for Efficacy Analyses in Oncology**

Author(s): William Wu, Xiaoxian Dai, Linda Gau

With the advancement of cancer care and the need to develop new therapies for cancer, oncology clinical trials have become more and more popular. Currently, most of the efficacy analyses in oncology clinical trials are represented by tables and listings. For patients with multiple visits, it is not convenient to read and understand the information presented in the tables and listings. In this paper, we introduce a SAS® macro developed in SAS 9.3 to create patient profile graphs in a multi-page PDF file using the SAS Output Delivery System (ODS) for efficacy analyses. On each page of the PDF file, there are two graphs. In addition, supporting information is added as listings to the graphs using the SAS Annotate Facility. Efficacy data for patients with chronic lymphocytic leukemia (CLL) are used to generate the graphics output using the developed SAS macro. Compared to the tables and listings, the graphs are more intuitive and much easier to understand.

---

Paper DG03

Tuesday: 9:30 - 9:50 a.m.

Room: Sapphire L

### **Developing Graphical Standards: A Collaborative, Cross-Functional Approach**

Author(s): Mayur Uttarwar, Murali Kanakenahalli

"A picture is worth a thousand words:" a quote proven by the fact that graphs are increasingly used in reporting clinical trial data. Conventionally, SAS® Graph procedures in conjunction with annotation datasets have been used to create graphs, but recently SAS® Graph Template Language (GTL) is gaining ground. GTL is very powerful in its ability to allow the building of complex and intricate graphs using a structured approach. However, the key factor in leveraging the full potential of GTL and maximizing program efficiency is to establish graphical standards. This paper addresses the need for graphical standards, showcases a collaborative approach with cross-functional partners to identify & standardize common graph attributes, and provides one style template & one graph template that can support multiple graph types.

Paper DG05

Monday: 9:00 - 9:50 a.m.

Room: Sapphire L

**JMP® Visual Analytics®**

Author(s): Charlie Shipp

For beginners and more experienced, we review the continuing merging of statistics and graphics. Statistical graphics is the forte of JMP software and the JMP team pioneered the way for SAS Visual Analytics®. Moving forward into Version 11, JMP has easy navigation, a graph builder, advanced algorithms in 85 platforms, and robust statistics you can trust. We also celebrate JMP success and discuss evolving and emerging developments. The 2013 JMP Discovery Summit presented new Design of Experiments additions and other additions to JMP 11. New also is the entry of the 12th platform for experimental design, the Definitive Screening Design, adding selected midpoints. Many of the views will be shown in the presentation, resulting in a lively discussion at the end!

---

Paper DG06

Monday: 11:45 - 12:05 p.m.

Room: Sapphire L

**Want to Conquer the Fear of Annotation? Start Using Note Statement.**

Author(s): Madhuri Aswale

The capabilities of #BYVAL in plotting graphs are exceptional for graphical display and can be made even more efficient and effective when used with the NOTE statement. Unlike TITLE and FOOTNOTE, NOTE is not frequently used in SAS code for creating and modifying graphs. The use of both NOTE and #BYVAL cohesively allows the programmer to write better code and to embellish graphs with essential information. The purpose of this paper is to explore the effective use of the NOTE statement to add text to maps, plots, charts, and text slides, resulting in better control of content, appearance, and placement of the text, including color, size, font, and alignment, in addition to allowing for underline and box drawing around text and straight lines on the output.

---

Paper DG07

Monday: 10:45 - 11:35 a.m.

Room: Sapphire L

**Swimmer Plot: Tell a Graphical Story of Your Time to Response Data Using PROC SGPLOT**

Author(s): Stacey D. Phillips

ODS Statistical Graphics (SG) procedures are making complex and sophisticated graphics easier to create with every new release of SAS. Through the use of color, line-types, symbols and annotations we can tell a complicated graphical story of our data in one glance. This paper will demonstrate how to create a swimmer plot using PROC SGPLOT that shows multiple pieces of tumor response information for individual subjects in an oncology study. Specifically, the swimmer plot will show total time to tumor response, whether the response was complete or partial, when the response started/ended, censoring information and the current disease stage of the subject. The paper will demonstrate the step-by-step process of creating a swimmer plot from basic PROC SGPLOT statements through more complex annotations using SAS version 9.3. There will also be some discussion on generating similar plots using SAS version 9.2 as well as Graph Template Language (GTL).

---

Paper DG08

Monday: 4:30 - 5:20 p.m.

Room: Sapphire L

**Napoleon Plot**

Author(s): Kriss Harris

Do you want to produce a very useful plot? Okay, do you want to produce a plot that for each subject shows the number of treatment cycles, the number of days on treatment, the doses that were received, whether the subject has discontinued treatment, and the cohort the subject is in? This paper will demonstrate how to do the above in version SAS® 9.3 and after.

---

Paper DG09

Tuesday: 9:00 - 9:20 a.m.

Room: Sapphire L

**Clinical Data Dashboards for Centralized Monitoring Using SAS® and Microsoft® SharePoint®**

Author(s): Jagan Mohan Achi

In recent guidance on risk-based monitoring practices, the Food and Drug Administration identified options for monitoring the data quality of a clinical investigation. Centralized monitoring is a practice in which clinical data management, statisticians, and/or clinical monitors<sup>1</sup> review the data on an ongoing basis and can easily identify the accuracy, completeness, and integrity of data. In this paper, we would like to describe the methods of developing an in-house tool using SAS® BASE (SAS® Institute, Cary, NC), GRAPH procedures (SAS® Institute, Cary, NC), and Microsoft® SharePoint® (Microsoft® Corporation, Redmond, WA) to achieve this goal. Centralized monitoring may decrease the cost of a clinical investigation and ensure identification of problems with the trial early on by maximizing the value of electronic data capture systems. Clinical dashboards created using SAS® BASE and GRAPH in conjunction with Microsoft® SharePoint® may provide cost savings, because most pharmaceutical/medical device companies frequently utilize these systems and no additional technology or human resources are required. Key Words: Risk Based Monitoring, Clinical Dashboards, SAS GRAPH

---

Paper DG10

Monday: 10:15 - 10:35 a.m.

Room: Sapphire L

**ODS EPUB: SAS® Output at Hand**

Author(s): Erica Goodrich, Daniel Sturgeon

A new addition to SAS® 9.4 is the Output Delivery System (ODS) Electronic-Publications (EPUB) destination. ODS EPUB can be used to create SAS® output files for various e-readers on both smart phones (e.g. Samsung Galaxy, Apple iPhone) and tablets (e.g. Apple iPad, and Amazon Kindle). We will discuss how to create SAS® output from varying reporting and graphical procedures alongside free flowing text and ODS EPUB customization options available to create easy to use e-publication files. Some PROCs may be discussed at moderate levels of complexity, however ODS EPUB topics will be discussed at an introductory level.

---

Paper DG12

Monday: 2:15 - 3:05 p.m.

Room: Sapphire L

**Automate the Process of Image Recognizing a Scatter Plot: an Application of a Non-parametric Statistical Method in Capturing Data from Graphical Output**

Author(s): Zhaojie Wang

A fundamental method of pharmaceutical research is to compare the indication between a new drug and a competitor drug. When a legacy graphical output is the only resource to access the competitor drug, researchers have to capture tabular data from it. Previously, a method was introduced to digitalize a graphical output into coordinates of a set of pixels. Then, the graphical output of a competitor drug can be relayed on the graph of the new drug for a comparison purpose. It may satisfy the need of reverse-engineering a curve plot, which present the trend, rather than individual spots in a scatter plot. To recognize a scatter plot correctly, it is necessary to identify the data presented by each individual spot as precisely as possible. Because an individual spot, even with the smallest size and the simplest pattern, can be composed with multiple pixels, it can be challenging to identify the corresponding data automatically from the coordinates of a set of pixels. In this paper, a non-parametric statistical method is introduced to facilitate recognizing a scatter plot. Windows SAS 9.2 is used to implement this non-parametric method automatically. It processes the coordinate data of all the pixels on the scatter plot to identify the data presented by each individual spot. The data obtained can be used for further analysis in the comparison. This automation strategy can improve not only the efficiency but also the accuracy of tabular data capture. Example plots and SAS codes are presented to illustrate this approach.

Paper DG13

Tuesday: 8:00 - 8:50 a.m.

Room: Sapphire L

### **I Am Legend**

Author(s): Kriss Harris

Have you ever produced a legend on a plot that was taking up too much space, hence making the actual graph too small? Have you ever removed a legend because it was taking up too much space? Have you ever wanted to just produce a legend? Have you ever wondered that there must be a more efficient way of producing a legend then using the exact same legend on every BY variable of your output? This paper will demonstrate solutions to the above problems using Graph Template Language (GTL) in SAS® 9.2, in particular using the SERIES, VECTOR and SCATTERPLOT statements.

---

Paper DG14-SAS

Monday: 1:15 - 2:05 p.m.

Room: Sapphire L

### **Up Your Game with Graph Template Language Layouts**

Author(s): Sanjay Matange

You have built the simple bar chart and mastered the art of layering multiple plot statements to create complex graphs like the survival plot using the SGPLOT procedure. You know all about how to use plot statements creatively to get what you need and how to customize the axes to achieve the look and feel you want. Now it's time to up your game and step into the realm of the "Graphics Wizard." Behold the magical powers of Graph Template Language layouts! Here you will learn the esoteric art of creating complex multi-cell graphs using LAYOUT LATTICE. This is the incantation that gives you the power to build complex, multi-cell graphs like the forest plot, stock plots with multiple indicators like MACD and Stochastic, Adverse Events by Relative Risk graphs, and more. If you ever wondered how the Diagnostics panel in the REG procedure was built, this paper is for you. Be warned, this is not the realm for the faint of heart!

---

Paper DG15-SAS

Tuesday: 10:15 - 10:35 a.m.

Room: Sapphire L

### **Quick Introduction to ODS DOCUMENT**

Author(s): Cynthia Zender

No abstract available.

---

Paper DS-PANEL

Tuesday: 11:15 - 12:05 p.m.

Room: Sapphire A

### **Panel Discussion: ADaM Implementation**

Moderator: Nancy Brucken

Many aspects of analysis dataset design are very study-specific. For that reason, items such as specific variables to include and the number of analysis datasets required for a given study are not defined in Version 1.0 of the ADaM Implementation Guide. Join us for an informal discussion of ADaM implementation details, and bring your own ADaM data set questions for our panel to address.

## Data Standards

---

Paper DS01

Monday: 1:15 - 2:05 p.m.

Room: Sapphire A

### **Discover Define.xml**

Author(s): Mark Wheeldon

Using CDISC standards in your day-to-day job can be complex and, maintaining Define can be a time consuming task. Formedix CEO, Mark Wheeldon will address frequently asked questions about Define-XML at PharmaSUG 2014. Mark will discuss the practical uses for Define in the course of an end-to-end clinical trial, a guide to implementation and an overview of what's new in Define 2.0. You will learn about: Creating and re-using Proprietary, CDISC SDTM and ADaM dataset libraries with Define Define-XML: The Myths and the Realities. Your Clinical Trials Automated. Everywhere. Study Start-up: Define-XML aided CRF design & specification process optimization Study Conduct and Analysis: Automated dataset validation with Define-XML The role of Define-XML in Legacy and Proprietary EDC dataset conversions Define-XML 2.0 enhancements: What do they mean to you? Define-XML 2.0 enhancements: Technical Deep Dive.

---

Paper DS02

Wednesday: 8:00 - 8:20 a.m.

Room: Sapphire A

### **Forging New SDTM Standards for In-Vitro Diagnostic (IVD) Devices: A Use-Case**

Author(s): Carey Smoak, Mansi Singh, Smitha Krishnamurthy, Sy Truong

How does a new data standard get established for medical devices? Data standards for medical devices have made good progress recently with the development of seven new SDTM domains specifically intended for medical device submissions. These seven new domains address the requisite domains to capture the data that is unique to medical devices because medical device data can be distinct and different from pharmaceutical and biotechnology data. These seven medical device domains were intended to capture data that is commonly collected across various types of devices. Currently, in SDTM for drugs, there is an on-going effort to develop therapeutic specific standards (e.g., Alzheimer's, Parkinson's, etc.). Similarly, within medical devices there is a need to develop standards for various types of devices. This paper addresses one such need to design domains specifically for In-Vitro Diagnostic (IVD) devices which are different from other medical devices (e.g., implantable devices). This paper will present a use-case for IVD devices. The project was undertaken at Roche Molecular Systems by a team that identified data used in IVD studies which can be generalized and implemented as an additional standard for IVD devices. The results are refinements to existing domains and creation of new domains along with variables that follow the standard established by CDISC. The goal of this paper and the team is to have these new standards be used in establishing the next set of SDTM and ADaM data models in support of IVD devices.

Paper DS03

Wednesday: 10:15 - 11:05 a.m.

Room: Sapphire A

### **Considerations in Creating SDTM Trial Design Datasets**

Author(s): Jerry Salyers, Fred Wood, Richard Lewis, Kim Minkalis

Many sponsors are now submitting clinical trials data to the FDA in the format of the CDISC SDTM. The Trial Design Model (TDM) datasets can be especially challenging because, in most cases, they are being created retrospectively from the protocol, and cannot be created from electronic data. From the most recent Common Data Standards Issues document, it is clear that FDA is placing a greater emphasis on incorporating the trial design datasets into any SDTM based submission. This presentation will discuss some of the considerations and challenges in creating the TDM datasets, using case studies of both relatively simple and more complex trials. We will highlight a number of the pitfalls and misconceptions that are commonly seen when sponsors and their vendors attempt to create the TDM datasets for the first time. Included will be practical advice on which datasets should be created first, which datasets drive the creation of others, and how the trial-level and subject-level datasets relate to each other. The presentation will conclude with a list of resources for TDM dataset creation.

---

Paper DS04

Tuesday: 9:00 - 9:50 a.m.

Room: Sapphire A

### **Considerations in the Submission of Exposure Data in SDTM-Based Datasets**

Author(s): Fred Wood, Jerry Salyers, Richard Lewis

The submission of data regarding the subjects' exposure to a study treatment is critical in assessing its safety and efficacy. While more and more sponsors are committing resources to submit SDTM-based datasets, our experience in legacy-data conversion has revealed that many studies don't collect sufficient data to get a reliable assessment of actual exposure. Even when such data have been collected, many sponsors are unsure of how to properly represent that data in the SDTMIG Exposure (EX) and Exposure as Collected (EC; new to SDTMIG v3.2) domains. This paper will discuss methods for representing exposure data, as well as some of the challenges sponsors may face in converting data to be consistent with the SDTMIG.

---

Paper DS05

Tuesday: 3:30 - 4:20 p.m.

Room: Sapphire A

### **Data Standards Development for Therapeutic Areas: A Focus on SDTM-Based Datasets**

Author(s): Fred Wood, Diane Wold, Rhonda Facile, Wayne Kubick

In recent years, there has been an increasing focus on developing data standards for therapeutic areas (TAs). A number of Therapeutic-Area User Guides (TAUGs) have been published via the collaboration of a number of standards-development organizations as well as the FDA. This presentation will provide an overview the TA development structure (CDISC, C-Path, CFAST, TransCelerate Biopharma, Inc.), the TAUGs, and some of the new SDTM-based domains and concepts. The relationship between domains in the TAUGs and versions of the SDTMIG will also be discussed. Finally, highlights for some of the domains new to the SDTMIG in v3.2 (e.g., Death Details, Microscopic Findings, Morphology, Procedures, and Skin Response), some of the provisional domains in TAUGs (e.g., Respiratory Measurements, Procedure Agents, and Meals), and the Disease Milestones concept will be presented.

---

Paper DS06

Monday: 2:15 - 2:35 p.m.

Room: Sapphire A

### **Referencing Medical Device Data in Standard SDTM domains**

Author(s): Timothy Bullock, Ramkumar Krishnamurthy

The seven new SDTM device domains capture the data describing patient exposure as well as all the details defining and identifying each device used in a medical device study. In order to provide continuity of device data across the entire SDTM model, some device-related data may also have to reside outside of the device domains. Examples of this are the AE domain which can contain information on device-related AEs and the PR domain which may need to accommodate a large amount of detailed information on device-related procedures such as implantation. We will discuss the integration of device data in the overall SDTM model and provide examples of methods for including device-related information in Findings domains outside of the device SDTM using the Findings About construct.

---

Paper DS07

Tuesday: 4:30 - 5:20 p.m.

Room: Sapphire A

### **Applying ADaM BDS Standards to Therapeutic Area Ophthalmology**

Author(s): Songhui Zhu

Most of the ADaM datasets are in BDS data structure. So statistical programmers spend a lot of time implementing BDS datasets. In this paper, the author illustrates how to implement ADaM-compliant BDS datasets when facing some complex situations in ophthalmology studies. The situations discussed are: 1) implementation of PARAM/PARAMCD for ophthalmology measurements; 2) implementation PARCATy for ophthalmology measurements; 3) implementation of by-visit baseline; 4) implementation of by-time-point baseline; 5) implementation of multiple baselines; 6) implementation of lab data using two sets of units; 7) implementation of LOCF.

---

Paper DS08

Monday: 9:00 - 9:50 a.m.

Room: Sapphire A

### **Challenges of Processing Questionnaire Data from Collection to SDTM to ADaM and Solutions using SAS®**

Author(s): Karin Lapann, Terek Peterson

Often in a clinical trial, measures are needed to describe pain, discomfort, or physical constraints which are visible but not measurable through lab tests or other vital signs. In these cases, researchers turn to questionnaires to provide documentation of improvement or statistically meaningful change in support of safety and efficacy hypotheses. For example, in studies (i.e. Parkinson's) where pain or depression are serious non-motor symptoms of the disease, these questionnaires provide primary endpoints for analysis. Questionnaire data presents unique challenges in both collection and analysis in the world of CDISC standards. The questions are usually aggregated into scale scores, as the underlying questions by themselves provide little additional usefulness. The SAS system is a powerful tool for extraction of the raw data from the collection databases and transposition of columns into a basic data structure in SDTM which is vertical. The data is then processed further as per the instructions in the Statistical Analysis Plan (SAP). This involves translation of the originally collected values into sums, and the values of some questions need to be reversed. Missing values can be computed as means of the remaining questions. These scores are then saved as new rows in the ADaM (analysis-ready) datasets. This paper describes the types of questionnaires, how data collection takes place, the basic CDISC rules for storing raw data in SDTM, and how to create analysis datasets with derived records using ADaM standards; while maintaining traceability to the original question.

**An ADaM Interim Dataset for Time-to-Event Analysis Needs**

Author(s): Kim Minkalis, Sandra Minjoe

The Clinical Data Interchange Standards Consortium (CDISC) Analysis Data Model (ADaM) Implementation Guide (IG) version 1.0 and the appendix document titled "The ADaM Basic Data Structure for Time-to-Event Analyses" each provide guidance for how to set up a dataset for producing a time-to-event (TTE) analysis. In practice, a single TTE analysis dataset is often used for analysis of many different events and censoring times. In fact, this TTE analysis dataset is often one of the most complicated created for a study. One of the biggest issues with TTE analyses is that there can be many different dates to consider for both the event and/or the censor. Some of these dates can be options for many different analyses - for example, date of death is the event in survival analysis, but can also be a censor date in time-to-response analysis. To make TTE analysis more clear, we've adopted the process of compiling an interim BDS (Basic Data Structure) dataset with all potential dates to be used in the TTE analyses. This paper explains our process and gives examples of how an interim dataset can be used to add traceability and understanding to a complex analysis.

---

**Developing ADaM Specifications to Embrace Define-XML 2 Requirements**

Author(s): Lin Yan

The CDISC Define-XML 2.0.0 standard was released in March 2013. The new standard has quite a few new features and requirements. In order to adopt the new standards, sponsors need make changes accordingly in the ADaM specification files, which are usually used as programming guidance and as the basis of the generation of ADaM define-XML as well. This paper will discuss the new features and requirements in the Define-XML 2.0.0 that have direct impact to ADaM specifications. Also, this paper will use examples to illustrate what changes/enhancements in the ADaM specifications are needed to meet the requirements specified in Define-XML 2.0.0.

---

**A Guide to the ADaM Basic Data Structure for Dataset Designers**

Author(s): Michelle Barrick, John Troxell

The Clinical Data Interchange Standards Consortium (CDISC) Analysis Data Model (ADaM) Implementation Guide (ADaMIG), published in 2009, describes the many components of a very powerful and flexible analysis dataset structure called the Basic Data Structure (BDS), and provides some rules and examples. The BDS is unique among CDISC data structures in the flexibility it provides for the addition of various kinds of derived rows to meet analysis needs. A companion CDISC document, "CDISC Analysis Data Model (ADaM) Examples in Commonly Used Statistical Methods," published in 2011, provides more in-depth examples of ADaM data and metadata solutions in particular scenarios. However, neither of the two documents provides a holistic explanation of the BDS that describes how the structurally-important variables and kinds of observed and derived rows function together. In this paper, the authors define categories of observed and derived rows. These definitions underpin a unified explanation of the BDS that provides an understanding of how the various kinds of rows and the structural BDS variables interact. Such knowledge is essential in order to design the appropriate solution for each data scenario and analysis need.

---

**Effective Use of Metadata in Analysis Reporting**

Author(s): Jeffrey Abolafia

Many organizations are effectively using metadata for the creation and validation of clinical and analysis databases. However the use of metadata for analysis reporting lags behind that of database production. While many sponsors have submitted Define.xml to document ADaM databases, very few of these define files have included analysis results metadata. The recent CDISC pilot demonstrated that results level metadata adds significant value to a regulatory submission. As the CDISC ADaM model has matured as an analysis dataset standard, the ADaM standard combined with results related metadata can be utilized to facilitate producing displays and statistical analysis and easily extended to generate the results portion of the Define.xml file. This presentation will examine how the ADaM standard used in conjunction with results level metadata can be used to both generate statistical reports more efficiently and substantially add value and traceability to the Define file.

**How Valued is Value Level Metadata?**

Author(s): Shelley Dunn

One of the challenges of implementing SDTM and ADaM is the vertical structure of some of the data and how variables can be described which are dependent on the test code (xxTESTCD) in SDTM or Parameter (PARAMCD) in ADaM. What criteria provide the best practice for determining when to use Value Level Metadata? The CDISC Define-XML Specification Version 2.0 document provides some general, albeit vague, rules for when to provide Value Level Metadata. "Value Level Metadata should be provided when there is a need to describe differing metadata attributes for subsets of cells within a column." "Value Level Metadata should be applied when it provides information useful for interpreting study data. It need not be applied in all cases." "It is left to the discretion of the creator when it is useful to provide Value Level Metadata and when it is not." The overriding message is that there are few requirements for what variables require this metadata and most of the criteria are based on the subjective notion of providing useful information. With requirements open to interpretation there are many correct ways to apply this metadata. What is considered useful to one stakeholder may or may not be useful to another. This presentation will use experience from a range of projects to look at why, when, and how to define Value Level Metadata balancing the amount of effort it takes to define this information with its value to stakeholders.

**CDISC Electronic Submission**

Author(s): Kevin Lee

FDA signed FDA Safety and Innovation Act (FDASIA) into law on July 9th, 2012 and announced Prescription Drug user Fee Act (PDUFA) V. And, FDA strongly recommended CDISC as electronic submission formats. The paper will introduce FDASIA, PDUFA V and Data Standard Strategy from FDA. Then, the paper will discuss what it means to programmers who prepare FDA submission. The paper will introduce where programmers can find FDA electronic submission guidelines and CDISC guidelines such as eCTD(electronic common technical document) specification, SDTM implementation guideline, ADaM implementation guideline and etc. First, the paper will provide the brief introduction of regulatory electronic submission such as its methods, five modules in CTD especially m5, technical deficiencies in submission and etc. And, the paper will discuss what programmers need to prepare for the submission according to FDA and CDISC guidelines for CSR, Protocol, SAP, SDTM annotated eCRF, SDTM datasets, ADaM datasets, ADaM datasets SAS programs and Define.xml. And secondly, the paper will discuss how programmers can prepare the submitted materials - length, naming conventions and file formats of electronic files. For examples, SAS data sets should be submitted as SAS transport file formats and SAS programs should be submitted as text format. Finally, the paper will discuss the latest FDA concerns and issues about the electronic submission such as the size of SAS data sets, the length of character variables in SAS datasets, CDISC compliance checks and etc.

Paper DS15

Tuesday: 1:15 - 2:05 p.m.

Room: Sapphire A

**A Road Map to Successful CDISC ADaM Submission to FDA: Guidelines, Best Practices & Case Studies.**

Author(s): Vikash Jain, Sandra Minjoe

Submitting a filing to the FDA (Food and Drug Administration) using Clinical Data Interchange Standards Consortium (CDISC) data standards has become a norm in the pharmaceutical industry in recent years. This standard has been strongly encouraged by agency to help expedite their review process, plus it also helps the sponsors and service providers have a means of efficient collaboration using these common industry standards. This paper elaborates on the following fundamental and core components to be considered for the ADaM (Analysis Data Model) piece of submission: 1) Use of ADaM or ADaM-like data 2) Checking for ADaM compliance 3) Use of Define and other Metadata We also present handful of case studies based on real time CDISC submission project experiences while collaboration with our sponsors.

---

Paper DS16

Tuesday: 2:45 - 3:05 p.m.

Room: Sapphire A

**OpenCDISC Validator Implementation: a Complex Multiple Stakeholder Process**

Author(s): Terek Peterson, Gareth Adams

The embracing of data standards by FDA following the vision of CDISC of reduced review time of drug applications opened the door for the creation of several tools to ensure conformance to standards. Tools like SAS PROC CDISC, WebSDM ", and SAS Clinical Standards Toolkit helped industry ensure compliance to the CDISC standards defined as SDTM, ADaM and define.xml. With the introduction of a free conformance engine OpenCDISC Validator, the possibilities of less confusion and more synergies across Sponsors, CROs and FDA was possible. However, the authors would argue the use of this tool has not achieved that goal and has created complex processes between stakeholders that include clinical, data management, programming, sponsor, and FDA; each group having different understanding of the conformance reports. Confounding any implementation are multiple versions OpenCDISC, SDTM, ADaM, and sometime contradicting FDA documentation. The way out of this confusion is with the implementation of good procedures, communication, and training. This paper will start with an example of waste where a clear process did not exist. It will provide examples of OpenCDISC checks that need to be managed early in the data lifecycle via edit checks to ensure fewer OpenCDISC warnings/errors. Communication and education needs to be in place for non-technical study team members so they can make informed decisions around the output. The paper provides processes to help control duplication of effort at different time points of a clinical trial. Budget considerations will be presented. Discussion and demonstration of example SAS® code will be provided.

---

Paper DS17

Monday: 11:15 - 11:35 a.m.

Room: Sapphire A

**Update: Development of White Papers and Standard Scripts for Analysis and Programming**

Author(s): Nancy Brucken, Michael Carniello, Mary Nilsson, Hanming Tu

A PhUSE Computational Science Symposium (CSS) Working Group is creating white papers outlining safety, demographics, disposition and medications analysis and reporting for clinical trials and regulatory submissions. An online platform for sharing code has also been created, making these standards easy to implement. This paper provides an update on the progress made in these efforts.

Paper DS18

Monday: 11:45 - 12:05 p.m.

Room: Sapphire A

### **An Alternative Way to Create Define.XML for ADaM with SAS Macro Automation**

Author(s): Yiwen Li

Define.XML for ADaM is required for most FDA submissions, as it describes the structure and contents of the ADaM data. It includes five sections: Data Metadata, Variable Metadata, Value Level Metadata, Computational Algorithm, and Controlled Terminology. Previously programmers used to create it with more support from filling out many Excel sheets as input. This paper provides a Linux SAS based simple method to extract almost all needed information from ADaM data with coding-logic-info-capture to create the whole Define.XML output. The only other importing source besides SAS is one Excel tab which stores description for variables derived with more complicated logic. It has been successfully implemented to HIV Phase1 study and significantly increased the efficiency of creating Define.XML by at least 50%. Highlighted Strengths: 1. Build most columns on Define.XML from SAS Libraries 2. Build Origin/Comment columns with coding logic information capture

---

Paper DS19

Monday: 10:45 - 11:05 a.m.

Room: Sapphire A

### **SAS® as a Tool to Manage Growing SDTM+ Repository for Medical Device Studies**

Author(s): Julia Yang

When we have a substantial number of medical device studies in many different therapeutic areas, it is desirable to have a common data repository to facilitate clinical data management, analysis and reports. Modified Study Data Tabulation Model plus (SDTM+) becomes the infrastructure for our clinical data. SDTM+, with some adaptations, followed the SDTM and the SDTM for Medical Devices (SDTM-MD) by Clinical Data Interchange Standards Consortium (CDISC). There are many challenges to map multiple studies in multiple therapeutic areas into one set of SDTM+ domains, including: 1) ensuring consistency across all studies; 2) incorporating new medical devices added periodically, quite possibly indefinitely; 3) making the SDTM+ database both scalable and stable; and 4) making sure the database is "self-explanatory" and source-traceable so new users do not need to refer to multiple documents. This paper summarizes what we learned so far. It discusses SAS macros developed to help map data to the SDTM+ , monitor SDTM+ consistency and check SDTM+ data integrity. Key Words: SAS, SAS Macro, CDISC, SDTM, SDTM-MD, Medical Device Data

---

Paper DS20-SAS

Wednesday: 8:30 - 9:20 a.m.

Room: Sapphire A

### **A How-To Guide for Extending Controlled Terminology Using SAS Clinical Data Integration**

Author(s): Melissa Martinez

SAS® Clinical Data Integration offers the ability to use additional controlled terminology beyond that provided by CDISC in order to standardize and validate data values. Terminology commonly used in this way includes MedDRA codes and the WHO Drug Dictionary, but even customized, company-specific terminology can be implemented. Users can also register newer versions of CDISC-published terminology that is not included in their version of SAS® Clinical Standards Toolkit. This paper fully describes the steps necessary to import extended controlled terminology into SAS Clinical Data Integration, use the terminology for compliance checks, and include or exclude the terminology from the define.xml document.

---

Paper DS21-SAS

Wednesday: 8:00 - 8:50 a.m.

Room: Sapphire D

### **Round Trip Ticket - Using the Define.xml file to Send and Receive your Study Specifications**

Author(s): Julie Maddox

Why can't you always get what you want? All you need to do is ask for it! Now that CDISC Data standards have matured, why not provide your project leads and external partners with the exact study specifications you require all wrapped up in a single CRT-DDS define.xml file. Code lists, domains, computational algorithms, and value-level metadata information are all captured in the CRT-DDS define.xml file. SAS Clinical Standards Toolkit provides a variety of macros to extract this study information from the define file and produce domain templates, format catalogs and metadata tables ready to be populated. SAS Drug Development provides a secure, web-based interface to a SAS programming environment with access to the SAS Clinical Standards Toolkit as well as the full power of the SAS analytics procedures and reporting features. SAS Drug Development also provides a secure repository to handle your organizations customized CDISC data standards and manage the study data in a compliant manner. This paper discusses how to round trip from a SDTM data standard, to a CRT-DDS define file and back out to a SDTM study using SAS Clinical Standards Toolkit within the SAS Drug Development SAS execution environment. Topics include: o managing CDISC standards o creating a CRT-DDS define file including controlled terminology, value-level metadata and computational algorithms o extracting metadata from a CRT-DDS define file o extracting code lists and domain templates from a CRT-DDS define file

---

Paper DS22-SAS

Tuesday: 10:15 - 11:05 a.m.

Room: Sapphire A

### **An Integrated platform to manage Clinical data, Metadata and Data Standards**

Author(s): Romain Rutten, Peter Wang, Sharon Trevoy

An integrated platform for clinical data, metadata and data standards management has been co-developed by Janssen, SAS Institute and Business and Decision Life Sciences to address critical business requirements of the global Data Management department. The comprehensive solution delivers business improvements in the following areas: " The preparation of clinical study specifications " Cross-study and cross-compound metadata comparison " Study validation for structure and content " Quality control and issue tracking " The use and maintenance of global and indication-specific libraries for SDTM and CDASH standards

---

Paper DS23-SAS

Tuesday: 8:00 - 8:50 a.m.

Room: Sapphire A

### **Creating Define-XML version 2 with the SAS® Clinical Standards Toolkit 1.6**

Author(s): Lex Jansen

In March 2013 the final version of the Define-XML 2.0.0 standard, formerly known as CRT-DDS (Case Report Tabulation Data Definition Specification) or "define.xml", as most people called it, was released by the CDISC XML Technologies team. Define-XML 2.0.0 is a major revision of the Define-XML standard for transmission of SDTM, SEND and ADaM metadata. Version 1.0.0 was released for implementation in February 2005. Define-XML has been a useful mechanism and critical component for providing Case Report Tabulations Data Definitions in an XML format for CDISC based electronic submissions to a regulatory authority such as the U.S. Food and Drug Administration (FDA). The Define-XML standard is based on the CDISC Operational Data Model (ODM). Version 1.0.0 was an extension to ODM version 1.2. The new Define-XML version takes full advantage of ODM 1.3.2. The Define-XML specification has been greatly improved with an increased clarity and reduced ambiguity. This presentation describes new features of Define-XML version 2, and will then describe how the SAS® Clinical Standards Toolkit can be used to create define.xml version 2 files.

## Healthcare Analytics

---

Paper HA01

Monday: 9:00 - 9:50 a.m.

Room: Sapphire H

### **Reporting Healthcare Data: Understanding Rates and Adjustments**

Author(s): Greg Nelson

In healthcare, we often express our analytics results as being "adjusted". For example, you may have read a study in which the authors reported the data as "age-adjusted" or "risk-adjusted." The concept of adjustment is widely used in program evaluation, comparing quality indicators across providers and systems, forecasting incidence rates and in cost-effectiveness research. In order to make reasonable comparisons across time, place or population, need to account for small sample sizes and case-mix variation - in other words - we need to level the playing field and account for differences in health status and accounts for uniqueness in a given population. If you are new to healthcare it may not be obvious what that really means to adjust the data in order to make comparisons. In this paper, we will explore the methods by which we control for potentially confounding variables in our data. We will do so through a series of examples from the healthcare literature in both primary care and health insurance. Included in this survey of methods, we will discuss the concepts of rates and how they can be adjusted for demographic strata (such as age, gender and race) as well as health risk factors such as case mix.

---

Paper HA02

Monday: 1:15 - 2:05 p.m.

Room: Sapphire H

### **Common and Comparative Incidence Indicators of Adverse Events for Well-defined Study Pools**

Author(s): John R Gerlach

Consider a large Integrated Safety Summary (ISS) study where the analysis of adverse events (AE) is based on well-defined study pools, such that each pool represents a unique collection of subjects. A subject might belong to more than one study pool, perhaps having a particular indication or some other criterion of clinical interest. Besides the standard adverse event reports, that is, a categorized event across treatment groups, it may be required to show incidence with respect to the degree of incidence, very common versus common, as well as in comparison between the study drug and placebo. The task of imputing so-called Common and Comparative incidence indicators becomes more involved since each indicator is determined by its own study pool. Moreover, the workload and its complexity have increased substantially. This paper explains the process of expanding the data set of adverse events (ADAE) and imputing these indicator variables for subsequent analysis.

---

Paper HA03

Monday: 2:15 - 3:05 p.m.

Room: Sapphire H

### **Survey of Population Risk Management Applications Using SAS®**

Author(s): Jack Shoemaker

The business of health insurance has always been to manage medical costs so that they don't exceed premium revenue. The PPACA legislation which is now in full force amplifies this basic business driver by imposing MLR thresholds and establishing other risk-bearing entities like ACOs and BPCI conveners. Monitoring and knowing about these patient populations will mean the difference between success and financial ruin. SAS® software provides several mechanisms for monitoring risk management including OLAP cubes, third-party solutions, and Visual Analytics. This paper surveys these SAS® solutions in the context of the population risk management problems now part of the healthcare landscape.

**Linking Healthcare Claims and Electronic Health Records (EHR) for Patient Management - Diabetes Case Study**

Author(s): Paul Labrec

Treo Solutions-now part of 3M Health Information Systems-conducted a pilot project to assess the feasibility of linking healthcare administrative claims data to an electronic health record (EHR) data extract to enhance patient case management activities. We linked one year of healthcare claims data (2012) to the equivalent year of medical record data abstracted from the EHR system of a large Midwest commercial insurer. The claims database identified 328,897 adult patients receiving services during 2012. Over 35,000 of these patients (10%) had a diabetes diagnosis. The clinical data set included 272,193 records on 61,532 patients in 2012 and included over 50 data elements. Measures identified in the EHR database included physical measures (the most common records), health history, health behaviors, radiologic and endoscopic tests, select prescription data and laboratory values. We abstracted a subset of EHR records for adults (ages 18-75) who had at least one diabetes-related test recommended by the National Quality Forum for use in this analysis. These tests include blood pressure, hemoglobin A1c, lowdensity lipoprotein, and retinal exams. From this combined database we calculated that the majority of patients with a diabetes diagnosis on claims had no diabetes test results for the study year. Furthermore, a small number of patients without a known diabetes diagnosis had at least one out-of-range diabetes test. We summarize the strengths and weaknesses of administrative claims versus EHR data for patient classification and compliance analyses, as well as methodological issues in combining claims and clinical databases. Planned follow-up analyses include medication fill rate calculations; cost of care predictions for various patient groups; and health outcomes analyses.

**%ME: A SAS Macro to Assess Measurement Equivalence for PRO (Patient-reported outcome) Measures**

Author(s): Qinlei Huang

Patient-reported outcomes (PRO) are the consequences of disease and/or its treatment as reported by patients. PRO measures are very important for new treatments, drugs, biological agents, and devices when used as effectiveness end points in clinical trials. Measurement equivalence is a function of the comparability of the psychometric properties of the data obtained via mixed modes (e.g., paper-pencil questionnaires, web-based questionnaires etc.); in paired data (e.g., patient reported outcome, parent reported outcome etc.); and/or at different time points. It is required to establish measurement equivalence prior to using a PRO measure for further statistical analysis and modeling. Multiple statistical methods have been developed to test measurement equivalence based on means (TOST, T-Test), variance (Levene's test), and correlations and agreements (Pearson Product Moment Correlation, Intra-class correlation, Weighted Kappa, and Spearman's Rho). A SAS Macro %ME is written to provide all the above mentioned statistics for measurement equivalence automatically at the same time. A sample with real data will be provided for illustration.

Keywords: Measurement equivalence, survey methodology

Paper HA06

Tuesday: 4:00 - 4:20 p.m.

Room: Sapphire D

**The Association of Morbid Obesity with Mortality and Coronary Revascularization among Patients with Acute Myocardial Infarction**

Author(s): Ashwini Erande

The aim of this study was to investigate the impact of morbid obesity (body Mass index  $\geq 40$  kg/m<sup>2</sup>) on in-hospital mortality and coronary revascularization outcomes in patients presenting with acute myocardial infarctions (AMI). The Nationwide Inpatient Sample of the Healthcare Cost and Utilization Project was used, and 413,673 Patients hospitalized with AMIs in 2009 were reviewed. Morbidly obese patients constituted 3.7% of all patients with AMIs. All the analyses were performed in SAS® 9.3. ARRAY Statements were used to create the morbid obesity variable based on the ICD9 codes from 24 "Diagnosis" data elements. The SAS procedures PROC SURVEYFREQ and PROC SURVEYLOGISTIC were used to perform bivariate and multivariate analyses for this sample survey data. The unadjusted and adjusted analyses performed using PROC SURVEYFREQ and PROC SURVEYLOGISTIC respectively, revealed that morbidly obese patients compared with those not morbidly obese were more likely to undergo any invasive coronary procedures when presenting with either ST-segment elevation myocardial infarction and also have a higher mortality rate. The SAS procedures used to analyze and summarize the data within this context are presented in this paper.

---

Paper HA07

Monday: 10:15 - 11:05 a.m.

Room: Sapphire H

**Using the SAS System as a bioinformatics tool: A macro that translates BLASTn results to populate a DNA sequence database table**

Author(s): Kevin Viel

Using the Basic Local Alignment Search Tool (BLAST) to align DNA sequences to a known reference sequence is a common task in genomics research. The results of the BLASTn (nucleotide) alignment can be translated into a database table of a relational database of a sequencing project. Appropriately annotating the reference sequences increases the utility of the resulting table. The National Center for Biotechnology Information (NCBI) of National Institutes of Health (NIH) provides the blast+ software package. The goals of this paper are to describe a SAS macro that uses the results of BLASTn alignments and Phred data to populate a table and to discuss some issues encountered in a sequencing project the generated data on over 60 million nucleotides.

---

Paper HA08

Tuesday: 3:30 - 3:50 p.m.

Room: Sapphire D

**PROC PHREG with RANDOM STATEMENT vs. PROC PHREG without RANDOM STATEMENT For Measuring Persistence Using Real World Evidence**

Author(s): Tracee Vinson-Sorrentino, Matthew Leichter, Yao Zhang

Market data correlation has always been a concern when designing a model to determine the overall persistence of a drug. Many times essential pharmaceutical elements are diluted in the data due to correlation of the patient demographics and physician treatment approaches. As a result the actual effects of a drug and adverse side effect factors become less certain. In some cases these factors can even be shown to be lacking in significance when, in actuality, they are significant. Logistical models that either ignore physician and patient demographic effects or assign a simple coefficient to them have considerably less power and specificity. Our objective is to show that the random effects model via the RANDOM statement in PROC PHREG is superior to the PROC PHREG standard model and will correct for this dilution. We found that random effects generalized models account for data correlation in patient groups within categorized demographics and by physicians. This allows for the removal of physician effects and patient demographic effects from the main drug factors by treating physician effects and patient demographics as random variables. Random effects generalized modeling removes improper data correlation, explains socioeconomic effects on patient behavior, and separates patient behavior from physician treatment when using a Cox-proportional hazards model to predict persistence.

---

**Using SAS® to Calculate and Compare Adjusted Relative Risks, Odds Ratios, and Hazard Ratios**

Author(s): Besa Smith, Tyler Smith

In the past decade, health outcomes research has gained in popularity as increasing focus has been given to improving patient outcomes. Whether upon refining screening strategies for earlier detection of disease, reducing readmission or nosocomial infection rates, improving patient satisfaction, or evaluating new patient therapies and treatments, the underlying focus is to prevent, control, and/or treat. The analysis of the occurrence of health outcome events is dependent upon differing risk sets among those with and without the event of interest and is often modeled using one of three approaches. Using regression approaches, we often see relative risk estimates, odds ratios, or hazards ratios presented after adjusting for a list of covariates that may be distorting our view. This paper will use SAS® to compare the process and results of a log-binomial regression, logistic regression, and Cox regression in the context of several covariates and including a temporal element. Discussion of why a researcher would use a certain approach in a specific situation will be discussed. Health outcome researchers strive to identify at-risk populations by providing quantitative evidence that allows for more informed decisions by practitioners and policy makers. This paper presents the code and results of three frequently used approaches in the evolving environment of health analytics.

**Estimating Medication Adherence Using a Patient-Mix Adjustment Method**

Author(s): Scott Leslie

The Centers for Medicaid and Medicare Services (CMS) and several national health care quality organizations regard medication adherence as a major attribute of quality of care. The preferred method of measuring medication adherence is the Proportion of Days Covered (PDC) by medication(s) over a specified review period. Although PDC can be calculated fairly easily from pharmacy claims using relatively little data elements, patient's medication adherence is most likely confounded by patient demographics and other measurable and immeasurable factors. This paper explores the use of a patient-mix adjustment method to account for patient's characteristics and previous medication history. Included is a description of a macro used to calculate PDC and estimate medication adherence.

---

## Hands-on-Training

---

Paper HT01

Monday: 10:15 - 11:45 a.m.

Room: Sapphire P

### **Programming With CLASS: Keeping Your Options Open**

Author(s): Art Carpenter

Many SAS® procedures utilize classification variables when they are processing the data. These variables control how the procedure forms groupings, summarizations, and analysis elements. For statistics procedures they are often used in the formation of the statistical model that is being analyzed. Classification variables may be explicitly specified with a CLASS statement, or they may be specified implicitly from their usage in the procedure. Because classification variables have such a heavy influence on the outcome of so many procedures, it is essential that the analyst have a good understanding of how classification variables are applied. Certainly there are a number of options (system and procedural) that affect how classification variables behave. While you may be aware of some of these options, a great many are new, and some of these new options and techniques are especially powerful. You really need to be open to learning how to program with CLASS.

---

Paper HT02-SAS

Monday: 3:30 - 5:00 p.m.

Room: Sapphire P

### **Practically Perfect Presentations**

Author(s): Cynthia Zender

PROC REPORT is a powerful reporting procedure, whose output can be "practically perfect" when you add ODS STYLE= overrides to your PROC REPORT code. This hands-on workshop will feature several PROC REPORT programs that produce default output for ODS HTML, RTF and PDF destinations. Workshop attendees will learn how to modify the defaults to change elements of PROC REPORT output, such as HEADER cells, DATA cells, SUMMARY cells and LINE output using ODS STYLE= overrides. In addition, attendees will learn how to apply conditional formatting at the column or cell level and at the row level using PROC FORMAT techniques and CALL DEFINE techniques. Other topics include: table attributes that control interior table lines and table borders, use of logos in output and producing "Page x of y" page numbering.

---

Paper HT03

Tuesday: 8:00 - 9:30 a.m.

Room: Sapphire P

### **Hands-On ADaM ADAE Development**

Author(s): Sandra Minjoe, Kim Minkalis

The Analysis Data Model (ADaM) Data Structure for Adverse Event Analysis<sup>1</sup> was released by the Clinical Data Interchange Standards Consortium (CDISC) ADaM team in May, 2012. This document is an appendix to the ADaM Implementation Guide (IG)<sup>1</sup>, and describes the standard structure of the analysis dataset used for most of our typical adverse event reporting needs. This hands-on training focuses on creating metadata for a typical adverse event (AE) dataset. Attendees will work with sample SDTM and ADaM data, finding information needed to create the results specified in a sample set of table mock-ups. Variable specifications, including coding algorithms, will be written. Some familiarity with SDTM data, AE reporting needs, SAS® data step programming, and Microsoft Excel® is expected. Attendees will also learn how to apply the data structure for similar analyses other than adverse events.

---

Paper HT04

Monday: 1:15 - 2:45 p.m.

Room: Sapphire P

### **A Hands-on Introduction to SAS Dictionary Tables**

Author(s): Peter Eberhardt

SAS maintains a wealth of information about the active SAS session, including information on libraries, tables, files and system options; this information is contained in the Dictionary Tables. Understanding and using these tables will help you build interactive and dynamic applications. Unfortunately, Dictionary Tables are often considered an 'Advanced' topic to SAS programmers. This paper and workshop will help novice and intermediate SAS programmers get started with their mastery of the Dictionary tables.

---

Paper HT05

Tuesday: 10:15 - 11:45 a.m.

Room: Sapphire P

### **Survival 101 - Just Learning to Survive**

Author(s): Leanne Goldstein, Rebecca Ottesen

Analysis of time to event data is common in biostatistics and epidemiology but can be extended to a variety of settings such as engineering, economics and even sociology. While the statistical methodology behind time to event analysis can be quite complex and difficult to understand, the basic survival analysis is fairly easy to conduct and interpret. This workshop is designed to provide an introduction to time to event analyses, survival analysis and assumptions, appropriate graphics, building multivariable models, and dealing with time dependent covariates. The emphasis will be on applied survival analysis for beginners in the health sciences setting.

---

Paper HT06

Tuesday: 1:15 - 2:45 p.m.

Room: Sapphire P

### **So You're Still Not Using PROC REPORT. Why Not?**

Author(s): Ray Pass, Daphne Ewing

Everyone who can spell SAS knows how to use PROC PRINT. Its primary use may be as a development tool to help in debugging a long multi-step program, or as a simple report generator when all that is really needed is a quick look at the data, or even a basic low-level finished report. However, if a report generation/information delivery tool with powerful formatting, summarizing and analysis features is called for, then PROC REPORT is the solution. PROC REPORT can provide the standard PROC PRINT functionality, but in addition, can easily perform many of the tasks that you would otherwise have to use the SORT, MEANS, FREQ and TABULATE procedures to accomplish. PROC REPORT is part of the Base SAS product, can run in both an interactive screen-painting mode or a batch mode, and should be the basic tool of choice when there is a need to produce powerful and productive reports from SAS data sets. This paper will present the basics of PROC REPORT (non-interactive mode) through a series of progressively more sophisticated examples of code and output.

---

Paper HT07

Tuesday: 3:30 - 5:00 p.m.

Room: Sapphire P

**SDTM, ADaM and define.xml with OpenCDISC®**

Author(s): Angela Ringelberg, Tracy Sherman

As programmers, many of us have spent hours reviewing SDTM/ADaM standards and implementation guides to generate "compliant" CDISC SAS data sets. However, there is an easier way to ensure compliance with CDISC standards, including SDTM, ADaM, Define.xml, and others. OpenCDISC® is an open source community which is focusing on creating frameworks and tools for the implementation and advancement of CDISC Standards. OpenCDISC® has created a CDISC Validator which will eliminate the need for individuals to develop their own custom processes in order to ensure that their CDISC models are compliant with CDISC standards. By taking common validation rules, OpenCDISC® has developed an open-source tool which is freely available and of commercial-quality to ensure data compliance with CDISC models such as SDTM, ADaM and Define.xml. The validation rules for each standard have been pooled into a CDISC Validation Rules Repository, providing users with a central listing. The listing is easy to use, modify and continue development. In this Hands-On Training, we are going to briefly describe a few of the key terms (SDTM, ADaM, Define.xml) and investigate the use of OpenCDISC Validator to perform the validation of SDTM, ADaM and define.xml.

---

Paper HT08-SAS

Wednesday: 8:00 - 9:30 a.m.

Room: Sapphire P

**Creating Multi-Sheet Microsoft Excel Workbooks with SAS®: The Basics and Beyond Part 1**

Author(s): Vince Delgobbo

This presentation explains how to use Base SAS®9 software to create multi-sheet Excel workbooks. You learn step-by-step techniques for quickly and easily creating attractive multi-sheet Excel workbooks that contain your SAS output using the ExcelXP ODS tagset. The techniques can be used regardless of the platform on which SAS software is installed. You can even use them on a mainframe! Creating and delivering your workbooks on-demand and in real time using SAS server technology is discussed. Although the title is similar to previous presentations by this author, this presentation contains new and revised material not previously presented.

---

## Industry Basics

---

Paper IB01

Monday: 9:00 - 9:50 a.m.

Room: Sapphire M

### **Challenges in Processing Clinical Lab Data**

Author(s): Alan Meier

Processing Clinical Laboratory data can be one of the most complex tasks during analysis data set creation, especially if the data are coming from local labs. This paper will examine some of these complexities within the CDISC SDTM LB data set structure, including: - Determination of how to identify and group the lab tests. This usually starts with assigning test codes and high-level groupings (hematology, chemistry, urinalysis, etc.). Deciding which codes should be used can be difficult, especially if the tests are not in the standard terminologies. There may also be the need to sub-group the tests (proteins, WBC differentials, etc.). - If central labs are not used, there may be a need to deal with tests that come in-house with different units. Where can conversion factors be found and how can an organization maintain and reuse them? - Safety labs often include "pesky" urine dipstick textual results (trace, 1+, etc.). What can be done with them? - Once the basic data set is processed, what criteria can be used to identify abnormalities and what are the complexities involved with them?

---

Paper IB02

Tuesday: 2:15 - 2:35 p.m.

Room: Sapphire D

### **TLF Validation Etiquette: What to Say, When to Say, How to Say, and Why to Say**

Author(s): Karen Walker

Peer review is the gold standard by where quality is achieved. It has been used to publish scientific journals for centuries. On the internet we see the effects of peer review every day. The peer review process has been mimicked in reality television shows that judge singers, dancers, and even for selecting a spouse. In this paper are the rules by where quality peer review can be done with success and bring forth further insights for SAS programs. Quality is achieved by using a checklist, cooperation, and support. Inherent within this process is the contrast of disparate ideas. That's the beauty of it. Flowing from those disparate ideas; contrast and compared, is proof that one idea is the best suited. Persons involved in this process must agree that the overall quest to discover the best is a collaboration effort. They have to be open to equally capable expertise among one and other, and truly share ideas. Those involved remain free of insult, "bullyism" and "come up isum". Should those latent idiosyncrasies prevail&they will take control. When ego takes over, a programmer and their peer reviewer can find themselves doing everything but collaborating. As a result, the most efficient ideas are lost to stand off, and the quality process is made useless. This is why I wrote this paper.

---

Paper IB03

Wednesday: 9:00 - 9:20 a.m.

Room: Sapphire D

### **Common Variables in Adverse Event and Exposure Analysis datasets specific for Oncology Study Trials**

Author(s): Hari Namboodiri

Cancer remains the second most common cause of death in the US with approximately 1 million new cases reported every year. According to the American Cancer Society, Cancer accounts for nearly 1 of every 4 deaths in the US. Cancer is also becoming prevalent in the developing world and estimated that over 21 million people will have cancer by 2030. As a result, more and more treatments are entering the market and sponsors are initiating new therapies in their Clinical Trials. Since Cancer is the general name for a group of more than 100 diseases, the initiation of Clinical Trial can be different for different types of Cancer. To add to this, a trial can further investigate whether the study drug is for a treatment, prevention or for screening and whether they can look at different types of cancers with the same study drug. To make things further complicated, sponsors have added another level of complexity over the last decade or so in their Study design to include "Biomarkers". Although, CDISC, has introduced a Tabulation model (IG 3.1.3), The ADaM IG for oncology is still not available. In this paper, I would present a list of specific variables for Adverse Events and Exposure analysis datasets which are currently not included in current ADaM Implementation guides and are necessary

for all Oncology specific analysis datasets. A brief overview of common Oncology specific analysis variables and their importance in study trials in specific types of cancer therapies will also be discussed.

---

Paper IB04

Monday: 10:15 - 11:05 a.m.

Room: Sapphire M

**Cover the Basics, Tool for structuring data checking with SAS**

Author(s): Ole Zester

Data Cleaning and checking are an essential part of the Statistical programmer's tasks. Therefore, many of us develop simple checks for these tasks. This paper presents a program package for an easier way to structure and over viewing data checks. The program is easy to use and it's very easy to implement new checks. If you have a common data structure you can use this to cover some of the basics, and you can use your time to search for the specials. If your data are changed dynamically i.e. not final, this program includes a way of remembering which issues you already have found. The program also includes a dynamic report which gives both an overview (how many findings) and the some more details (which finding).

---

Paper IB05

Wednesday: 9:30 - 9:50 a.m.

Room: Sapphire D

**Attain 100% Confidence in Your 95% Confidence Interval**

Author(s): Indu Nair, Binal Patel

A very common mistake in the calculation of a confidence interval occurs when there are no qualifying subjects in a by group for a category being tested. It is very tempting to assume that the confidence interval will be missing when the count is zero, which is incorrect. This tends to get overlooked since the usual methods of calculating confidence intervals, such as using a simple PROC FREQ in SAS®, will not take care of the situation without manipulating the code further. This paper will present the different methods in PROC FREQ that allow you to calculate the confidence intervals and discuss which methods are more appropriate to use. This paper will also explain how to use a formula instead of PROC FREQ to calculate confidence intervals correctly and with confidence.

---

Paper IB06

Monday: 11:15 - 12:05 p.m.

Room: Sapphire M

**Good versus Better SDTM - Why "Good Enough" May No Longer Be Good Enough When It Comes to SDTM**

Author(s): Henry Winsor

While companies are finally making strong efforts to use and provide SDTM, mostly because of serious encouragement from the FDA, there seems to be some misunderstandings about why this work needs to be done. Although originally strictly intended to replace paper Case Report Form Tabulations with data that is electronically accessible, SDTM is being used - and abused -- with other purposes in mind. Problems arise when people forget what SDTM is for and tailor their implementations for something other than easing the Reviewer's task. The authors provide a gentle reminder of SDTM's real purpose in the FDA submission world and some suggestions of how and why to maximize your company's benefit from having SDTM data sets available for use, preferably long before a CSR or eCTD is complete.

Paper IB07

Wednesday: 10:15 - 10:35 a.m.

Room: Sapphire D

**From "just shells" to a detailed specification document for tables, listings and figures.**

Author(s): Supriya Dalvi

We are assigned a new study. We go through the protocol, the statistical analysis plan, the mock up shells and start programming. The basic annotations for the mock up shells is been done. The programmer uses this as a specification document for generating the outputs; similarly the validator uses it for QC of the outputs. There are some differences in understanding between the two as the shells are not "clear enough" leading to discussions amongst them , also involving the statistician at several times .At the time of Statistical review, it is observed there are inconsistencies in the layout of the reports, date formats, visit orders, treatment group information, common derivations etc& The result: a lot of rework on correcting these issues, loss of time, confusions and a question mark on the quality of the reports. Could this have been avoided if the mock up shells were given some more attention than what they got? May be Yes&How? Let's try to understand&

---

Paper IB09

Tuesday: 2:45 - 3:05 p.m.

Room: Sapphire D

**Clinical Study Report Review: Statistician's Approach**

Author(s): Amita Dalvi

A clinical study report (CSR) is one of many types of regulatory documents that comprise a marketing application for a drug, biologic, or device. A study statistician is a co-author of CSR which is a descriptive account of a single clinical trial accompanied by tables, listings, and figures (TLFs) displaying all study data and results. The study statistician works closely with the medical writer to ensure clarity and accuracy in conveying statistical findings and interpreting results, and addresses any statistical questions. This paper will discuss a study statistician's responsibilities while preparing and reviewing the clinical study report. The checklist for study statistician will make clinical study report review relatively easy.

---

Paper IB10-SAS

Tuesday: 1:15 - 2:05 p.m.

Room: Sapphire D

**Clinical Trial Data Transparency: Seeing is Believing**

Author(s): Janet Stuelpner

In 2012, the idea of pharmaceutical companies providing their proprietary clinical trials data as a public resource would have met with skepticism and disbelief. Fast-forward to 2013 and 2014, and pharmaceutical companies are racing to market with implemented solutions that not only enable them to share their own company trial data, but enables that data to be combined with data from other pharmaceutical companies. Whether it's to stay ahead of emerging guidance from the European Medicines Agency, or to simply document the integrity of their research programs, one thing is clear - clinical trial data transparency is a critical topic to understand in 2014.

---

## JMP

---

Paper JMP-PANEL                      Tuesday: 2:15 - 3:05 p.m.                      Room: Sapphire H

### **Panel Discussion: JMP and JMP Training**

Moderator: Charlie Shipp

A panel of JMP invited speakers and JMP users will each present a viewpoint on JMP software, JMP usage, and JMP user and management training. We will discuss the importance of management visibility, user group possibilities, and JMP training within your enterprise and in the community. You, the audience, will be an important part of the lively discussion that follows.

---

Paper JP01-SAS                      Tuesday: 1:15 - 2:05 p.m.                      Room: Sapphire H

### **Risk-Based Monitoring of Clinical Trials Using JMP® Clinical**

Author(s): Kelci Miclaus

Guidelines from the International Conference on Harmonisation (ICH) suggest that clinical trial data should be actively monitored to ensure data quality. Traditional interpretation of this guidance has often led to 100 percent source data verification (SDV) of respective case report forms through on-site monitoring. Such monitoring activities can also identify deficiencies in site training and uncover fraudulent behavior. However, such extensive on-site review is time-consuming, expensive and, as is true for any manual effort, limited in scope and prone to error. In contrast, risk-based monitoring makes use of central computerized review of clinical trial data and site metrics to determine whether sites should receive more extensive quality review through on-site monitoring visits. We demonstrate a risk-based monitoring solution within JMP® Clinical to assess clinical trial data quality. Further, we describe a suite of tools used for identifying potentially fraudulent data at clinical sites. Data from a clinical trial of patients who experienced an aneurysmal subarachnoid hemorrhage provide illustration.

---

## **Management & Support**

---

Paper MS-PANEL

Tuesday: 10:15 - 11:05 a.m.

Room: Sapphire M

### **Panel Discussion: Today's Marketplace for Statistical Programmers & Consultants**

Moderators: Jim Baker and Kent LeTourneau

Is the employment future and career paths for Statistical Programmers still bright? Experts in recruiting and placing Statistical Programmers, Biostatisticians, and Data Managers will engage in a lively and informational discussion about the current market place. What trends are being experienced in the market place involving: Salaries, Supply & Demand, Outsourcing, Skill Sets, Work from Home, Hiring, and Expected Experience. The panelist will be in a unique position of understanding and knowing how both the candidate expectations match with the market expectations. Questions and answers will be intertwined within the topical discussions to create a dynamic interaction with the audience.

---

Paper MS01

Wednesday: 8:00 - 8:20 a.m.

Room: Sapphire M

### **A New Trend in the industry - Partnership between CROs and Pharma. Do we know how to work in this new relationship?**

Author(s): Kevin Lee

There is a new trend in the working relationships between CROs and Drug companies. The relationships evolve from the traditional transactional model to a partnership in the recent years. The partnership in the pharmaceutical industry is a new adventure to both CROs and Drug companies, and it also impacts programmers. The partnership will bring the new challenges and opportunities to programmers. We will work with not only our colleagues, but also employees from other vendors and clients. The partnership working environments will require the new set of rules and behaviors. The paper will discuss what kind of challenges and opportunities we will encounter and how we will resolve and succeed in the partnership working environments. The paper will specifically discuss the role of a project leader in CRO who are working directly with the clients and other vendors in the partnership environments. It will emphasize how important the culture is in the partnership environments and how the project leader helps to build the right cultures in the partnership environments.

---

Paper MS02

Tuesday: 3:30 - 4:20 p.m.

Room: Sapphire M

### **Building Better Programming Teams with Situational Exposure Training**

Author(s): Elizabeth Reinbolt, Steve Kirby

A programmer's life in the broadly collaborative, timeline driven world of clinical research requires the ability to handle a wide variety of situations. While it is (relatively) simple to find good materials to teach a programmer specific SAS skills such as ODS or macro processing, the most challenging issues programmers (and programming teams) often face are not a lack of technical skills but difficulty thriving in unfamiliar situations. Although there are a host of companies ready to share general strategies for common workplace issues such as time or stress management, the authors suggest that, just as sports teams mimic game situations in practice, one useful way to help programmers and programming teams thrive in new situations is to expose them to those situations through role playing within the group.

---

Paper MS03

Tuesday: 8:00 - 8:50 a.m.

Room: Sapphire M

**Distance Management: how to lead a team of SAS users who sit half a world away**

Author(s): Max Cherny

This focus of the paper will be the author's experience of managing a large group of SAS users (both programmers and statisticians) based in India. The paper will describe advantages and challenges of managing such a group as well as various project management techniques to guide the projects to successful completion

---

Paper MS04

Wednesday: 9:00 - 9:20 a.m.

Room: Sapphire M

**Was Dorothy Right; Is There No Place Like Home?**

Author(s): Kjersten Offenbecker

Many companies within the pharmaceutical industry have adopted the practice of allowing their SAS programming staff to work from home either part-time or full-time. To many programmers this trend offers several benefits including reduced travel time, less disruption in family life, improved work-life balance and more flexible hours. For management and the company as a whole the benefits include cost savings, increased productivity, improved employee motivation, employee retention and improved employee satisfaction. But there are drawbacks to working from home for both the employee and the company including loss of focus, poor working environment as well as loss of connection and/or collaboration with other team members. This paper will explore the pros and cons of working from home from the perspective of someone who works from home full-time and successfully manages employees who work both as telecommuters and participate in more traditional working situations.

---

Paper MS05

Tuesday: 1:15 - 2:05 p.m.

Room: Sapphire M

**How To Win Friends and Influence People - A Programmer's Perspective on Effective Human Relationships.**

Author(s): Priscilla Gathoni

Dealing with people has become a task and an art that every person has to master in the work and home environment. This paper explores 15 different ways that a programmer can use to win friends and influence people. It displays the steps leading to a positive, warm, and enthusiastic balanced work and life environment. The ability to think and to do things in their order of importance is a key ingredient for a successful career growth. Programmers who want to grow beyond just programming should enhance their people skills in order to move up to the management level. However, for this to be a reality a programmer must have good technical skills, possess the ability to arouse enthusiasm among peers, and is able to assume leadership. It is the programmer that embraces non-judgment, non-resistance, and non-attachment as the core mantras that will succeed in the complex and high paced work environment that we are in. Avoiding arguments, being a good listener, respecting the other person's point of view, and recalling people's names will increase your earning power and ability to influence people to your way of thinking. The ability to enjoy your work, be friendly, and be enthusiastic tends to bring you goodwill. This eventually leads to creating good relationships in the office and the power to influence those around you in a positive way.

---

Paper MS07

Tuesday: 9:00 - 9:20 a.m.

Room: Sapphire M

### **The Fourth Lie - False Resumes**

Author(s): Ernest Pineda

"There are three types of lies - lies, damn lies, and statistics - Benjamin Disraeli". The fourth lie - false resumes. Twelve years ago we were surprised to see an increase in the number of resumes which exaggerated the sender's breadth of experience and work history. And now, twelve years later, we are still receiving resumes which are suspect. Unfortunately, not all false resumes are exposed and individuals are getting jobs that may endanger a clinical trial. Over time we've discovered that false resumes can be vetted before they hit the hiring managers desk. In this presentation Ernest Pineda, President of the Gerard Group will discuss measures his company has taken to validate the accuracy of resumes using pre-planned questions, inexpensive background checks, and industry knowledge. He will share experiences, observations, policies and tools that have helped his firm expose under qualified, and falsely represented candidates.

---

Paper MS08

Tuesday: 2:15 - 2:35 p.m.

Room: Sapphire M

### **Demystify "Ten years of pharma programming experience required" - What hiring managers actually look for**

Author(s): Peng Yang

Pharmaceutical SAS programmer job descriptions often times specify a certain number years of experience as a requirement. But what exactly does it mean to have 5 or 10 years of experience? Does one person's 10-year count the same as another person's 10-year? In this paper, we will discuss the relationship between industry experience length and some key qualifications for a successful pharma SAS programmer. Some of the qualification is not quantifiable and elusive to explain, thus the number of years is used as one of the simplified criteria in job posting. Discussion around the qualification can help recruiters conduct more effective screening and it can also enable candidates to write more specific and informative resumes for their job search.

---

Paper MS09

Tuesday: 2:45 - 3:05 p.m.

Room: Sapphire M

### **Monitoring Quality, Time and Costs of Clinical Trial Programming Projects using SAS®**

Author(s): Jagan Mohan Achi

Effective management of the Quality, Time and Costs (QTC) of programming activities entails a collective collaboration between statisticians and programming teams. In order to manage a portfolio of studies, there is always a necessity of having a close monitoring of the progress by the management. This ensures that the projects are delivered with a great control over QTC. To manage the projects, data is collected for tracking (programming trackers), operation (resource projections), finance (baseline costs and continuing costs) and quality (issues trackers). Project dashboards can be built with SAS GRAPH® using the operational data with risk metrics built around the project progress. This paper will describe how SAS® can help in monitoring project progress in a simple and yet, elegant way in order to keep your projects on track and have a better understanding of the implications of QTC on a real time basis. Keywords: Project Management, Dashboards, Quality, Time, Cost

---

Paper MS10

Wednesday: 8:30 - 8:50 a.m.

Room: Sapphire M

### **KAIZEN**

Author(s): Usha Kumar

In an ever-growing, demanding and changing world, we need to understand that the pursuit for excellence is endless. There needs to be a process of continuous improvement. This being the basis of my paper, I would elaborate on certain principles from the management theory that can be practiced at individual level. The process of continuous improvement at individual level when practiced by many in an organization develops a CULTURE. QUALITY becomes a habit; CULTURE builds a successful and healthy workplace. I present KAIZEN and other principles that can help implement KAIZEN as a self-improving technique. We will see some principles that help us implement KAIZEN. Kaizen involves setting standards and then continually improving on those standards. I would take examples from the clinical industry here. We are the best people who know our strengths and weaknesses; what we do right and what can be further improved. If we look back at our routine task and really try to assess, we would discover many more areas where we could do much better.

---

Paper MS11

Tuesday: 11:15 - 11:35 a.m.

Room: Sapphire M

### **A Guide for Front-Line Managers to Retain Talented Statistical SAS® Programmers**

Author(s): R. Mouly Satyavarapu

SAS® programmers have a variety of functional and technical skills, coding preferences, personality types, and communication styles. And those programmers, who aspire to make the career move to a manager role, should be aware of the skills that comprise a manager's skill set. I am presuming that a SAS® programmer who has been working towards a management role has gained the initial management title and is aware of the fundamental skills (grouped under): Project Management, Personnel Management, Communication, Negotiation, Team Motivation, Conflict Resolution, and Financial Management, which is referred in the paper - manager's skill set (La Brec 2010) [1]. In this paper, I will provide tips for front-line managers or project team leaders to - build a strong and productive Statistical SAS® programming team; focus on developing the team; recognize and reward the effort by each team member; and retain the programming team's talent. Additionally, in each of the above mentioned areas, I will list the required skill and emphasize few competencies within each of them, which could be considered as the key elements in working towards the penultimate goal of each leader to retain best performers/ programmers and their talent.

---

Paper MS12

Tuesday: 9:30 - 9:50 a.m.

Room: Sapphire M

### **Recruiting for Retention**

Author(s): Kathy Greer, John Reilly

Recruiting in the hopes of identifying and retaining highly qualified staff becomes more difficult each year. Since DataCeutics' is a Full Service Provider, our staff is placed in many highly visible positions throughout our clients. Following a process that eliminates applicants that cannot function independently and produce high level deliverables is critical. We follow a proven process to hire employees. The process includes determination of communication skills, following directions, abstract thinking, understanding of clinical statistical programming by a verbal discussion and programming, and lastly a background check. The applicant speaks with at least three staff who determines the possibility of a fit within our company. This paper will go into detail on each step followed during the hiring process.

**When is Validation Valid?**

Author(s): Janet Stuelpner

Validation is a topic that is discussed repeatedly each time a program is written for data preparation and analysis of clinical trials. The current methodologies are time consuming, resource intensive and not always as accurate as we would like them to be. Use of newer technology, like code generation tools such as SAS® Clinical Data Integration and SAS® Enterprise Guide, can make validation easier. Improvement in accuracy of the validation effort as well as making the process easier and more efficient is a common goal. Will a code generation tool provide less variation in programming? Will the use of industry standards and adaptation of metadata allow for more review and less dual programming? This paper will show different approaches to validation of data preparation and reporting outputs from clinical trials.

---

## Posters

---

Paper PO01                      On Display Monday and Tuesday                      Room: Sapphire Level

### **A User Friendly Tool to Facilitate the Data Integration Process**

Author(s): Yang Wang, Boxun Zhang

For any clinical program, meta-analysis, including Development Safety Update Report (DSUR), Integrated Summary of Safety (ISS) and Integrated Summary of Efficacy (ISE), integrating data is a challenging process and is often labor intensive & error prone. The challenges arise primarily due to the following reasons: (1) different study designs and data standards, (2) evolving industry standards including CDISC®, MedDRA and WHODrug, (3) different data collection methods across vendors. In order to ensure all the variables can be integrated for analysis, we built a user friendly tool to dynamically display the differences in variable attributes and values across multiple clinical studies. This paper, with examples and programming components, shows how this tool can be used to facilitate data integration by identifying and resolving issues resulting due to variable inconsistencies.

---

Paper PO02                      On Display Monday and Tuesday                      Room: Sapphire Level

### **Survival Analysis Procedures and New Developments Using SAS**

Author(s): Jane Lu, David Shen

A common feature of survival data is the presence of censoring and non-normality. It is inappropriate to analyze survival data using conventional statistical methods such as linear regression or logistic regression, because of the characteristics of survival data. When censoring, survival time can't be considered as a continuous variable. Linear regression compares mean time-to-event between groups, it cannot properly handle censored data. Logistic method compares proportion of events between groups using odds ratio, but the differences in the timing of event occurrence are not considered. And analyzing the probability of survival as a dichotomous variable by Chi-square test fails to account for the non-comparability in unequal survival times between subjects. This paper provides an overview of survival analysis and describes its principles and applications. SAS examples are provided to illustrate the procedures of LIFEREG, LIFETEST, PHREG and QUANTLIFE for survival analysis. New developments for survival analysis including time-dependent covariates, recurrent events, quantile regression will also be discussed.

---

Paper PO04                      On Display Monday and Tuesday                      Room: Sapphire Level

### **CDISC Mapping and Supplemental Qualifiers**

Author(s): Arun Raj Vidhyadharan, Sunil Jairath

Mapping of datasets from sponsor defined data structure, otherwise known as clinical data management (CDM) data structure, to CDISC SDTM structure can be one of most trickiest and complex programming situations. There are several methods devised by companies for this purpose. Some use SAS Mapping tools, some use tools based on SAS, VB and Excel while some use just SAS programs. Irrespective of the technique used, the basic fundamentals of mapping remain the same. This paper covers the various factors to be considered while mapping, certain unique scenarios one can encounter and possible solutions to them. This paper also covers the creation of SUPPLEMENTAL QUALIFIERS and custom domains and talks about a powerful SAS procedure to validate your SDTM datasets.

---

Paper PO07

On Display Monday and Tuesday

Room: Sapphire Level

### **SAS Can Automatically Provide GTL Templates for Graphics in Three Ways**

Author(s): David Shen, Li Zhang, Ben Adeyi, Dapeng Zhang

User-defined GTL templates are needed for the complicated statistical graphics which can not be obtained directly from the SG procedures. However, GTL is a relatively new language to many SAS users and it encompasses a large amount of syntax, statements and options. Actually there are three ways SAS may automatically provide the GTL templates. This presentation will show you how we can leverage this SAS feature to obtain the desired GTL templates, and then tailor these templates to create the customized statistical graphics, without writing GTL codes from scratch.

---

Paper PO08

On Display Monday and Tuesday

Room: Sapphire Level

### **Design of Experiments (DOE) Using JMP®**

Author(s): Charlie Shipp

JMP has provided some of the best design of experiment software for years. The JMP team continues the tradition of providing state-of-the-art DOE support. In addition to the full range of classical and modern design of experiment approaches, JMP provides a template for Custom Design for specific requirements. The other choices include: Screening Design; Response Surface Design; Choice Design; Accelerated Life Test Design; Nonlinear Design; Space Filling Design; Full Factorial Design; Taguchi Arrays; Mixture Design; and Augmented Design. Further, sample size and power plots are available. We give an introduction to these methods followed by a few examples with factors.

---

Paper PO09

On Display Monday and Tuesday

Room: Sapphire Level

### **Bad Dates: How to Find True Love with Partial Dates**

Author(s): Namrata Pokhrel

This poster will discuss the difficulties encountered while trying to impute for partial dates for a specific ISS/ISE study with particularly dirty data. This study contained multiple types of partial dates within a data set. Character dates were in a smorgasbord of DATE9, YYMMDD10, YYMMDD8 formats. Some dates had dashes, some had UNK or UK or nothing at all to represent the missing piece of the date. There were also many partial dates that were simply invalid. The poster will discuss the lessons learned while trying to impute for these partial dates and will provide recommendations for future SAS programmers.

---

Paper PO10

On Display Monday and Tuesday

Room: Sapphire Level

### **Switching from PC SAS to SAS Enterprise Guide**

Author(s): Cindy (Zhengxin) Yang

As more and more organizations adapt to SAS Enterprise Guide, switching smoothly from PC SAS to SAS Enterprise Guide and reducing the learning curve become important. The tips from work experiences can help maintaining productivity efficiency and revealing the powerful capabilities of SAS Enterprise Guide during the switch. This paper presents below tips of using SAS Enterprise Guide from firsthand experiences: Setting up working environment by choosing preferred options in tool manual is a good start of programming. Using work space layout, 'send to' option and other tips to utilize MS applications as additional tools can improve work efficiency; Grouping programs as projects will make it very easy to work with multiple programs and even multiple

projects at the same time. These applicable ways of using SAS Enterprise Guide will ensure a fast mastering of SAS Enterprise Guide for both new and experienced SAS programmers, particularly the programmers working in Pharmaceutical industries.

---

Paper PO11                      On Display Monday and Tuesday                      Room: Sapphire Level

**Guidelines for Protecting Your Computer, Network and Data from Malware Threats**

Author(s): Ryan Paul Lafler, Kirk Paul Lafler

Malware, sometimes referred to as malicious software, represent software threats engineered to damage computer systems, networks and data without the knowledge of the owner using the system. All users are increasingly becoming more prone to malware attacks and need to have strategies and a set of guidelines to help them get the most out of their antivirus software. This poster highlights a classification of malware threats, the types of computer threats, detection strategies, removal methods, and provides guidelines for the protection of essential assets.

---

Paper PO12                      On Display Monday and Tuesday                      Room: Sapphire Level

**Process and Tools for Assessing Compliance with Standard Operating Procedures**

Author(s): Ginger Redner, Eunice Ndungu

In our efforts to develop drugs and vaccines that improve and save lives, we must maintain the integrity of patient data. This is accomplished by ensuring compliance with SOPs (Standard Operating Procedures) for developing and validating programs used to generate output for the Clinical Study Report (CSR) and submission deliverables. We have developed a suite of macros to help study teams ensure their deliverables are compliant with our SOPs. The main compliance macro identifies and reports specific areas of non-compliance. Another assesses the use of previously developed and fully validated standard macros. Using these macros greatly reduces the risk of error in reporting, decreases the number of SOP steps required by the project programmers to generate the deliverables, and increases efficiency by saving programming time. This paper discusses the macros and the process followed to ensure consistency and mitigate deviations.

---

Paper PO13                      On Display Monday and Tuesday                      Room: Sapphire Level

**Adopted Changes for SDTMIG v3.1.3 and 2013 OpenCDISC Upgrades**

Author(s): Yi Liu, Stephen Read

There have been several enhancements and upgrades to CDISC Study Data Tabulations Model (SDTM) data standards and associated implementation guides over recent years. August 2012 saw the release of SDTMIG v3.1.3 and adoption of SDTMIG v3.1.2 Amended 1 recommendations, in advance of the more recent and comprehensive SDTMIG v3.2 upgrade as released in December 2013. In support of these SDTM enhancements the OpenCDISC community released 2 new SDTM validators in 2013. OpenCDISC Version 1.4 released in March 2013 and OpenCDISC 1.4.1 released in September 2013. Adoption of these enhanced SDTM data standards and associated validation applications has led to the need for SAS programmers or data mapping/submission specialists to consider implementing significant revisions across SDTM related SAS programs, processes and applications. This paper will give a brief background and overview of some of the key SDTMIG and OpenCDISC updates and offer some potential guidance in support of their implementation with real case examples on potential SDTM programming enhancements for SAS programmers or mapping specialist working with the newer OpenCDISC validators, focusing primarily on some welcome improvements based on the following SDTM recommendations: 1. Interpretations on Required/Expected/Permissible variables including; a) Inclusion of Study day variables such as STDY and ENDY across the majority of SDTM domains b) The addition of EPOCH to all clinical subject-level observation domains 2. Overview of revisions to TS (Trail design) domain 3. Formatting and character variable length/size setting across all SDTM databases 4. Some recommendations on supporting documentation and associated SDTM reviewers guides

Paper PO14

On Display Monday and Tuesday

Room: Sapphire Level

**Route to SDTM Implementation in In-Vitro Diagnostic Industry: Simple or Twisted**

Author(s): Carey Smoak, Sofia Shamas, Chaitanya Chowdagam, Lim Dongkwan

SDTM implementation for In-Vitro Diagnostic(IVD) data found a guiding light when CDISC came up with seven new domains specific to the medical device industry. Although these domains are not tailored for IVD data, they serve as a good starting point for us to define SDTM+, which is an extension of the SDTM standards to accommodate the large variety of IVD data. Similarity across studies following SDTM domain structure presents opportunities for code standardization and re-usability. What this paper hopes to achieve is to take the reader on our journey and explain briefly the standards created, the new domains and how our data was adjusted across SDTM+ domains. With the main focus on mapping of data, specific case examples will be listed describing programmatic techniques used in deriving the new standards in a systematic way.

---

Paper PO15

On Display Monday and Tuesday

Room: Sapphire Level

**Evaluating the benefits of JMP® for SAS programmers**

Author(s): Xiaopeng Li, Chun Feng, Nancy Wang

JMP® is a user friendly software for creating figures in clinical research. For SAS programmers, JMP® also provides a convenient way to do quality control (QC) of listings, tables, and figures (TFLs). Additionally, Using JMP® for QC is an alternative way to verify statistical analysis results, a quick way to verify SAS code for statistical analysis tables, and a way to do QC documentation. This paper provides details about what supports and benefits JMP® can provide for SAS programmers.

---

Paper PO16

On Display Monday and Tuesday

Room: Sapphire Level

**Automation of ADaM Dataset Creation with a Retrospective, Prospective, and Pragmatic Process**

Author(s): Karin Lapann, Terek Peterson

In the CDISC Standards, analysis datasets using the standards (ADaM) hold a unique place in the end to end process and must be created with both a prospective and retrospective view of the entire clinical trial process. Analysis datasets must support the statistical analysis plan (SAP) by providing accurate efficacy and safety analyses. Companies must be pragmatic in deciding which processes can be automated by tools. Industry has tools to effectively transform data to the SDTM structure. These tools should be able build a large portion of the appropriate ADaM datasets with maximum efficiency. The burning question: Can ADaM datasets be built with a mapping tool just like SDTM? The theme of this poster is to describe how standards can aid in the process of automating analysis datasets, to give programmers and biostatisticians more time to focus on the science and unique analyses for new indications and treatments. Automated processes require the proper governance by sponsors internally, and through collaboration between Clinical Research Organizations and the Sponsors. The decisions made interpreting the standards need to be assimilated and documented using a Metadata Repository (MDR) so that rigorous and consistent implementation can be assured. The MDR provides consistent input to the processing of the data from collection to Tables, Figures and Listings (TFL's) which are then used by medical writers to create the Clinical Study Report (CSR).

Paper PO17

On Display Monday and Tuesday

Room: Sapphire Level

### **Healthcare Data Manipulation and Analytics Using SAS**

Author(s): Lumin Shen, Jane Lu

Increasing application of information technology in the healthcare delivery system helps healthcare industry gain valuable knowledge from data, and use this insight to recommend action or guide decision making, and improve quality of patient care and practice of medicine. There is more data available than ever before. How can it truly benefit patients, payers and healthcare providers? Analytics can help medical researchers exploit healthcare data to discover knowledge lying implicitly in individual patients' health records, physicians identify effective treatments and best practices, patients receive better and more affordable healthcare services, and healthcare insurance companies detect fraud and abuse. This article explores analytics applications in healthcare industry. It illustrates the process of data integration and exploration, and building of predictive models to find previously unknown patterns and trends. It also presents analytics applications in major healthcare areas.

---

Paper PO18

On Display Monday and Tuesday

Room: Sapphire Level

### **Compare Without Proc Compare.**

Author(s): Pavan Vemuri

In clinical trial programming, the quality of data is of paramount importance. To ensure the quality of data parallel programming is performed where the programmer and the validator program independently followed by comparison of the two results. When the comparison involves data sets, PROC COMPARE is widely used. The validation of data sets is a twofold process where a) the attributes of variables and number of observations in both the data sets is matched and then b) the values of observations themselves are compared. Often times when deriving analyses data sets like Exposure, Labs etc. or programming tables that involve a lot of conditions, trying to match the number of observations can become tricky. Although using PROC COMPARE with ID statement helps in identifying mismatches, the need to sort the data and in some cases identifying the unique ID variables themselves can become time consuming. This paper proposes a macro that helps in reducing the time taken to resolve the discrepancies on number of observations. The macro does not require the data to be sorted and shows mismatches as counts of observations grouped by a user identified variable. The idea is to give an overall pattern on the nature of mismatches. Once the number of observations is equal, comparing the data becomes easy and PROC COMPARE can be used on the final run.

---

Paper PO19

On Display Monday and Tuesday

Room: Sapphire Level

### **An Overview of REDCap, a secure web-based application for Electronic Data Capture**

Author(s): Kevin Viel

REDCap (Research Electronic Data Capture) is a secure web application for building and managing online surveys and databases. With the simplest use, REDCap automates the interactions with the SAS System, providing an option to manually download files, including SAS code to create a SAS datasets and format library. Involvement of the SAS System® in order of increasing sophistication includes using SAS to write the "data dictionary" or template for the electronic forms, using an API to access the projects, or to read the MySQL databases directly. This paper provides a very brief introduction to REDCap and a superficial overview of the using SAS System with REDCap.

---

Paper PO20

On Display Monday and Tuesday

Room: Sapphire Level

**A Parameterized SAS Macro to Select an Appropriate Covariance Structure in Repeated Measures Data Analysis Using PROC MIXED**

Author(s): Paul Nguyen, Charity Quick, Leela Aertker

In clinical trials we often encounter multiple measurements on a subject or experimental unit over a period of time. When analyzing longitudinal data using SAS PROC MIXED, it is critical to specify the covariance or correlation structure of the repeated measures. Unless stated a priori, one must determine the most appropriate covariance structure by comparing and contrasting the best fit of different covariance structures based on AIC, AICc, BIC, or -2 log likelihood. This paper will describe an efficient macro to determine the covariance structure for a model. The %COVSTRUC macro includes several user-specified parameters: goodness-of-fit statistic, model information, and whether a particular covariance structure has priority given convergence. The covariance structure with the best fit (i.e., lowest value of specified fit statistic) will be chosen and outputted to a macro variable, which can then be called into the final analysis model.

---

Paper PO22

On Display Monday and Tuesday

Room: Sapphire Level

**Tips for Finding Your Bugs Before QC Does**

Author(s): Beatriz Garcia

Statistical Programmers often have to work as both programmers and in validation or verification of programs during the course of a project. Sometimes after we deliver the program for validation, we receive many requests to change the results or update the code. Why does that happen? Maybe because we have not checked the results as carefully as we should. In this paper, if you work in UNIX environment, I want to show you some tips for reviewing your results before delivering to QC, so you can prevent many of those requests, avoid re-work and save time.

---

Paper PO24

On Display Monday and Tuesday

Room: Sapphire Level

**Create Excel TFLs Using the SAS Add-in**

Author(s): Jeffrey Tsao, Tony Chang

"Can you create this table, figure or listing in Excel?" We often receive such a request in our day to day work. Using Excel to create TFLs is a good method to rapidly generate simple TFLs using SAS add-in that will save time and programming resources. Product: SAS add-in 5.1 for MS office. Operating system: MS Windows. SAS version: metadata server SAS 9.2 or above in Windows or UNIX. Skill level: All users.

---

## Statistics & Pharmacokinetics

---

Paper SP01                      Monday: 3:30 - 3:50 p.m.                      Room: Sapphire H

### **Factor analysis of Scale for Assessment of Negative Symptoms using SAS Software**

Author(s): Ben Adeyi, David Shen

The objective is to study the psychometric properties of the SANS (Scale for the Assessment of Negative Symptoms) using factor analysis, and investigate the feasibility of shortening the current version of the SANS. SANS data were examined with exploratory factor analysis to identify the principle components of SANS. The results showed that the SANS consisted of three factors, suggested a short SANS with 3 response cluster options of 10 items. Confirmatory factor analysis demonstrated the reliability of the short version of SANS.

---

Paper SP02                      Monday: 4:00 - 4:20 p.m.                      Room: Sapphire H

### **Handling with Missing Data in Clinical Trials for Time-to-Event Variables**

Author(s): Giulia Tonini, Simona Scartoni, Angela Capriati

Missing data is often a major issue in clinical trials, especially when the outcome variables come from repeated assessments. In particular, time-to-event endpoints can be substantially affected by a too conservative treatment of missing data along the observation period. When neglected or not properly treated, missing data may bias the results, reduce power and lead to wrong study conclusions. The advantage of a more sophisticated statistical method versus the traditional clinical method, such as last observation carried forward (LOCF), is still under debate. We compare the two methods in a clinical study testing the efficacy of an anti-arrhythmic agent versus placebo on the time to Atrial Fibrillation (AF) recurrence, where the maintenance of normal heart rhythm or the occurrence of the AF event was to be daily evaluated by trans-telephonic ECG recorded by the patients. A Cox model is applied for the comparison between treatments. The dataset presents missing observations due to the fact that recording is missing or ECG is not assessable. Moreover a simulation is performed to provide an additional example. Both methods for handling missing data are applied. Multiple imputation in SAS uses PROC MI. We examined results and possible problems arising from the fact that PROC MI implements methods which are not suitable for this kind of data.

---

Paper SP03                      Monday: 5:00 - 5:20 p.m.                      Room: Sapphire H

### **Defining Non-Inferiority Margins for Skin Adhesion Studies**

Author(s): Marina Komaroff, Sailaja Bhaskar

The non-inferiority trials might be performed as opposed to, or in addition to the superiority trials. The aim in non-inferiority trials is to demonstrate that the test product is "not worse" than reference listed drug (RLD) or active control by more than the non-inferiority margin. The non-inferiority studies carry some weaknesses such as assay sensitivity, blinding, defining appropriate non-inferiority margins. Yet, the use of active control in non-inferiority trails might be the only choice to demonstrate the efficacy of a new product when the use of placebo arm is not ethical. Another example is transdermal products where it is recommended that the adhesion performance of the test patch should be compared to the adhesion performance of RLD. One of complexities of non-inferiority trials is to define an appropriate non-inferiority margin. The choice of such margin is usually based on historical data from the previous trials and assumes that the same effect will be present in the planned non-inferiority trial. Regulatory guidance documents recommend a non-inferiority test for comparing adhesion performance of transdermal test product compared to adhesion performance of RLD. This paper has a goal to review and compare the existing regulatory guidelines for the conducting non-

inferiority trials particularly to compare adhesion performance. The authors suggest the rationale for adhesion data analyses with the clinically meaningful choice of non-inferiority margin. The examples of adhesion data analysis for two simulated studies are provided. The authors are convinced that this novel paper can make non-inferiority trials comparing adhesion performance more clinically relevant.

---

Paper SP04

Tuesday: 8:00 - 8:50 a.m.

Room: Sapphire H

**%AIC\_LOGISTIC: A SAS Macro to Produce Sorted Information Criteria (AIC/BIC) List for PROC LOGISTIC for Model Selection**

Author(s): Qinlei Huang

Model selection is one of the fundamental questions in statistics. One of the most popular and widely used strategies is model selection based on information criteria, such as Akaike Information Criterion (AIC) and Sawa Bayesian Information Criterion (BIC). It considers both fit and complexity, and enables multiple models to be compared simultaneously. However, there is no existing SAS procedure to perform model selection automatically based on Information Criteria for PROC LOGISTIC, given a set of covariates. This paper provides a set of SAS macros to select a final model with the smallest value of AIC/BIC. Specifically, it will 1) produce a complete list of all possible model specifications given a set of covariates, 2) run models using SAS Procedure PROC LOGISTIC, 3) use SAS/ODS to report AICs and BICs produced by PROC LOGISTIC with model specifications, and 4) append all reports to create a sorted list of model specifications and corresponding AICs/BICs. Based on this list, analysts could find the best model among all the possible variable combinations. This paper includes the macro programming language, as well as an example of the macro call and output. Keywords: Model Selection, Information Criterion, SAS/ODS, SAS Macro

---

Paper SP05

Tuesday: 9:00 - 9:20 a.m.

Room: Sapphire H

**A SAS® Macro to address PK timing variables issues**

Author(s): Timothy Harrington

This paper discusses and lists a SAS® macro which addresses the challenge of evaluating timing variables and chronological sequencing for collected PK sampling and patient dosing data when sample and dose dates and times are not always recorded correctly on, or are omitted from, the case report form input. Examples of such datasets are PKS, POPPK, and NONMEM. This macro is intended for use with the input data in CDISC format for a single dose per visit, with a single pre-dose sample and following post dose samples.

---

Paper SP06

Tuesday: 9:30 - 9:50 a.m.

Room: Sapphire H

**A Mental Health and Risk Behavior Analysis of American Youth Using PROC FACTOR and SURVEYLOGISTIC**

Author(s): Deanna Schreiber-Gregory

The current study looks at recent health trends and behavior analyses of youth in America. Data used in this analysis was provided by the Center for Disease Control and Prevention and gathered using the Youth Risk Behavior Surveillance System (YRBSS). A factor analysis was performed to identify and define latent mental health and risk behavior variables. A series of logistic regression analyses were then performed using the risk behavior and demographic variables as potential contributing factors to each of the mental health variables. Mental health variables included disordered eating and depression/suicidal ideation data while the risk behavior variables included smoking, consumption of alcohol and drugs, violence, vehicle safety, and sexual behavior data. Implications derived from the results of this research are a primary focus of this study. Risks and benefits of using a factor analysis with logistic regression in social science research will also be discussed in depth. Results included reporting differences between the years of 1991 and 2011. All results are discussed in relation to current youth health trend issues. Data was analyzed using SAS® 9.3.

---

Paper SP07

Monday: 4:30 - 4:50 p.m.

Room: Sapphire H

### **A SAS Macro to Evaluate Balance after Propensity Score Matching**

Author(s): Erin Hulbert

Propensity score matching is a method used to reduce bias in observational studies by creating two populations that are similar (i.e., balanced) across a number of covariates using a match on only a single scalar, the propensity score. The matched samples are then treated as a quasi-experimental population, allowing for simplified analysis of study outcomes. Whenever a propensity match is performed, the balance between the two samples should be evaluated. Balance checking may be used to compare matches from multiple iterations of the propensity score model or from different matching algorithms and to provide information for any trade-offs between the closeness of the match and final sample size. Additionally, imbalances in the final matched sample should be kept in mind and possibly adjusted for when analyzing study outcomes. The published literature encourages using a variety of methods to evaluate balance. We have developed a SAS macro that will run a series of tests on the pre- and post-matched samples, following published guidelines. This macro evaluates balance independent of the methods used to create the propensity score or perform matching. Macro output includes information about the pre- and post-matched sample sizes, distribution of propensity scores, and the results of tests to compare covariates of interest, including calculating standard differences. The output is in RTF format so that it can easily reviewed by non-programmers. This SAS macro to evaluate balance with easy-to-read output allows for thorough testing of balance and contributes to a better final matched sample.

---

Paper SP08

Tuesday: 4:30 - 5:20 p.m.

Room: Sapphire H

### **Methodology for Non-Randomized Clinical Trials: Propensity Score Analysis**

Author(s): Dan Conroy

Randomized clinical trials serve as the gold standard for all research trials conducted within the pharmaceutical industry. When trials are not properly randomized, there is a potential for bias in all subsequent statistical analyses. When proper randomizations are not in place for a trial, methods exist that can help researchers draw valid conclusions. This paper summarizes and demonstrates some potential methods that can be used in this scenario. In particular, propensity score methodologies are presented and discussed with illustrative examples and SAS code. The examples and issues discussed herein were selected with the intent that researchers may find them informative, relevant, and applicable in a variety of scenarios, so that they may mimic and apply these methods in their own statistical analyses. The content is directed at users of SAS software with an intermediate understanding of standard statistical concepts and methodologies.

---

Paper SP09

Wednesday: 9:00 - 9:20 a.m.

Room: Sapphire H

### **Same Data, Separate MEANS - 'SORT' of Magic or Logic?**

Author(s): Naina Pandurangi, Seeja Shetty

Sample Mean is the most fundamental element of any statistical analysis. It is also considered to be one of the simplest descriptive statistic with a straightforward formula i.e. sum of individual sample values divided by the sample size. Quite reasonably, one would expect to get a unique mean value for a given set of results irrespective of the order of the individual data points. But does this notion hold true forever or can there be some infrequent cases to challenge this apparent assumption of uniqueness of mean for a pre-defined sample in SAS? Can we really get different mean values on the same set of data, in two different situations - once when data is sorted by some variable and second when it isn't? Sounds a bit absurd and unthinkable but yes, we are talking about more than one mean value on the same data in SAS! This paper discusses some peculiar and uncommon examples of data in SAS which have the capacity to yield more than value of sample mean when pulled in, in a sorted and un-sorted format.

Paper SP10

Wednesday: 10:15 - 10:35 a.m.

Room: Sapphire H

**Our Survival Confidence Intervals are not the Same!**

Author(s): David Franklin

"Our Survival Confidence Intervals are not the same!" This was a message I received from a client who was doing some checking of my statistics. After some research it was found that our results were "correct" insofar as the programming was correct but the internal calculations by the software was different. This paper looks into the calculation of the Kaplan-meier estimate and compares the different methods of how the quartiles are calculated and the formulas for the confidence intervals -- they are not only different among different software but also among different versions of SAS. Along the way we shall see some of the options that affect the calculations and see how the defaults have changed in recent releases of SAS. There will also be a macro that calculates some options that SAS does not do.

---

Paper SP13

Wednesday: 9:30 - 9:50 a.m.

Room: Sapphire H

**The Path Less Trodden - PROC FREQ for ODDS RATIO**

Author(s): Manjusha Gondil

In our clinical industry we conduct various types of simple/complex statistical analysis for data interpretation and Odds Ratio is one such extremely useful and simple to understand technique. Odds Ratio is used for making the critical decision about which treatment is benefiting the trial subjects in a clinical trial. In statistics, the odds of an event occurring is the probability of the event divided by the probability of an event not occurring. Historically we calculate odds ratio using complex stats procedure like PROC LOGISTICS, PROC GENMOD etc., but this paper gives us an idea about easier alternative of how to find odds ratio using simpler stats procedures like PROC FREQ. Programmers who do not have background of statistics will find this simplified method of finding odds ratio very easy to code and understand. This method allows anyone to program in a much easier way.

---

Paper SP14-SAS

Tuesday: 3:30 - 4:20 p.m.

Room: Sapphire H

**Customizing the Kaplan-Meier Survival Plot in PROC LIFETEST in the SAS/STAT® 13.1 Release**

Author(s): Warren Kuhfeld

If you are a medical, pharmaceutical, or life sciences researcher, you have probably analyzed time-to-event data (survival data). The LIFETEST procedure computes Kaplan-Meier estimates of the survivor functions and compares survival curves between groups of patients. You can use the Kaplan-Meier plot to display the number of subjects at risk, confidence limits, equal-precision bands, Hall-Wellner bands, and homogeneity test p-value. You can control the contents of the survival plot by specifying procedure options with PROC LIFETEST. When the procedure options are insufficient, you can modify the graph templates with SAS macros. PROC LIFETEST in the SAS/STAT® 13.1 release provides many new options for Kaplan-Meier plot modification, and the macros have been completely redone in this release in order to provide more power and flexibility than was found in previous releases. This paper provides examples of these new capabilities.

Paper SP15-SAS

Tuesday: 10:15 - 12:05 p.m.

Room: Sapphire H

### **Modeling Categorical Response Data**

Author(s): Maura Stokes

Logistic regression, generally used to model dichotomous response data, is one of the basic tools for a statistician. But what do you do when maximum likelihood estimation fails or your sample sizes are questionable? What happens when you have more than two response levels? And how do you handle counts? This tutorial briefly reviews logistic regression for dichotomous responses, and then illustrates alternative strategies for the dichotomous case and additional strategies such as the proportional odds model, generalized logit model, conditional logistic regression, and Poisson regression. The presentation is based on the third edition of the book *Categorical Data Analysis Using the SAS System* by Stokes, Davis and Koch (2012). An existing working knowledge of logistic regression is required for this tutorial to be fully beneficial.

---

Paper SP16

Wednesday: 8:00 - 8:50 a.m.

Room: Sapphire H

### **Automating Pharmaceutical Safety Surveillance process**

Author(s): Chandramouli Raghuram

Pharmaceutical companies invest huge amount of effort and cost in launching a drug into the market, keeping the two aspects in mind i.e. safety and efficacy. Post launching the drugs into the market, the organisations then need to monitor the adverse events from these drugs, and need to take the action accordingly. In this white paper few approaches are discussed for automating the post market safety surveillance processes in a cost effective manner. The pharmaceutical companies collect the adverse events data from various heterogeneous sources, and this collected data need to be analysed for the safety surveillance. Generally in post market safety surveillance processes, each Drug-Event case is processed record by record which is causing the exponentially rise of the records and finally leading to the high computational complexity and Analytics performances issues in the system. This paper will also focus on handling these Performances issues by bringing Hadoop environment into the Solution Implementation. The technical SAS solution to be discussed in this white paper will include components for Data Extraction and Transformation, Analysis, Reporting and Automated Processes for Signals detection & Investigation.

---

## **Techniques & Tutorials: Foundations**

---

Paper TT01

Wednesday: 8:00 - 8:50 a.m.

Room: Sapphire L

### **Modernizing Your Data Strategy: Understanding SAS Solutions for Data Integration, Data Quality, Data Governance and Master Data**

Author(s): Greg Nelson, Lisa Dodson

For over three decades, SAS has provided capabilities for beating your data into submission. In June of 2000, SAS acquired a company called DataFlux to add data quality capabilities to its portfolio. Recently, SAS folded Data Flux into the mother ship and with SAS 9.4, the SAS Enterprise Data Integration (and baby brother Data Integration) solutions were upgraded into a series of new bundles that still include the former DataFlux products, but those products have grown. These new bundles include data management, data governance, data quality and master data management and come in advanced and standard packaging. This paper will explore these offerings and help you understand what this means to both new and existing customers of the Data Integration and DataFlux products. We will break down the marketing jargon and give you real world scenarios of what customers are using today (pre-SAS 9.4) and walk you through what that might look like in the SAS 9.4 world. Each scenario will include what software is required, what each of the components do (features and functions) as well as the likely architectures that you may want to consider. Finally, for existing Data Integration customers, we will discuss implications for migrating to the new version and detail some of the functionality that may be new to your organization.

---

Paper TT02

Wednesday: 9:00 - 9:20 a.m.

Room: Sapphire L

### **Are You Missing Out? Working with Missing Values to Make the Most of What is not There**

Author(s): Art Carpenter

Everyone uses and works with missing values, however many SAS® programmers are unaware of the variety of tools, options, and techniques associated with using missing values. Did you know that there are 28 types of numeric missing values? Did you know that the numeric missing value (.) is neither the smallest or largest possible numeric missing value? Are you aware of the System options, DATA step functions, and DATA step routines that specifically deal with missing values? Do you understand how the macro null value is the same, and different from DATA step missing values? Are you aware that observations with missing classification variables may or may not be excluded from analyses depending on the procedure and various options? This paper explores various aspects of the world of missing values. The above questions and others are discussed. Learn more about missing values and make sure that you are not missing out.

---

Paper TT04

Monday: 1:45 - 2:05 p.m.

Room: Sapphire M

### **'V' for & Variable Information Functions to the Rescue**

Author(s): Richann Watson, Karl Miller

There are times when we need to use the attributes of a variable within a data set. Normally, this can be done with a simple CONTENTS procedure. The information can be viewed prior to programming and then hardcoded within the program or it can be saved to a data set that can be joined back to the main data set. If the attributes are hardcoded then what happens if the data set changes structure, then the program would need to be updated accordingly. If the information from PROC CONTENTS is saved and then joined with the main data set, then this would need to be done for all data sets that need to be processed. This is where knowing your 'V' functions can come in handy. The 'V' functions can be used to return the label, format, length, name, type and/or

value of a variable or a string within the data step. These functions can come quite in handy when you need to create summary statistics and if you need to perform an algorithm on a variable with a specific naming convention.

---

Paper TT05                      Tuesday: 4:30 - 5:20 p.m.                      Room: Sapphire L

**Principles of Writing Readable SQL**

Author(s): Ken Borowiak

PROC SQL is a commonly used data manipulation utility by SAS users. Despite its widespread use and logical components, there is no accepted standard for writing PROC SQL. The benefits of well-written code include ease of comprehension and code maintenance for both the original author and those who inherit it. This paper puts forth some principles of writing readable SQL in an objective manner. Each principle is followed by a suggested coding style. The ultimate goal is to elucidate the attributes of well-written PROC SQL code that users will implement to maximize comprehension by all.

---

Paper TT06                      Wednesday: 10:15 - 10:35 a.m.                      Room: Sapphire L

**Functioning at a Higher Level: Using SAS® Functions to Improve Your Code**

Author(s): Peter Eberhardt, Lucheng Shao

SAS provides many built in functions that will help you write cleaner, faster code. With each new release of SAS there are new functions added; in many cases these new functions are overlooked because we have developed coding habits to try to accomplish the same result as these functions. In this paper we will survey some functions we find useful. In addition we will touch on how you can turn your code habits into functions.

---

Paper TT08                      Tuesday: 2:15 - 3:05 p.m.                      Room: Sapphire L

**Investigating the Irregular: Using Perl Regular Expressions**

Author(s): Peter Eberhardt

A true detective needs the help of a small army of assistants to track down and apprehend the bad guys. Likewise, a good SAS® programmer will use a small army of functions to find and fix bad data. In this paper we will show how the small army of regular expressions in SAS can help you. The paper will first explain how regular expressions work, then show how they can be used with CDSIC.

---

Paper TT09                      Tuesday: 3:30 - 4:20 p.m.                      Room: Sapphire L

**Strategies and Techniques for Debugging SAS® Program Errors and Warnings**

Author(s): Kirk Paul Lafler

As a SAS® user, you've probably experienced first-hand more than your share of program code bugs, and realize that debugging SAS program errors and warnings can, at times, be a daunting task. This presentation explores the world of SAS errors and warnings, provides important information about syntax errors, input and output data sources, system-related default specifications, and logic scenarios specified in program code. Attendees learn how to apply effective techniques to better understand, identify, and fix errors and warnings, enabling program code to work as intended.

---

Paper TT10

Tuesday: 1:15 - 2:05 p.m.

Room: Sapphire L

---

**Strategies and Techniques for Getting the Most Out of Your Antivirus Software for SAS® Users**

Author(s): Ryan Paul Lafler, Kirk Paul Lafler

Malware, sometimes referred to as malicious software, represent software threats engineered to damage computer systems without the knowledge of the owner using the system. SAS® users are increasingly becoming more prone to malware attacks and need to have a set of guidelines to help them get the most out of their antivirus software. This presentation highlights the many different types of computer threats, classification approaches, detection strategies, and removal methods. Attendees learn what malware is; the types of malware including viruses, Trojans, rootkits, zombies, worms, spyware, adware, scareware, spam email, and denial of service (DOS) attacks; password protection and management strategies; software to detect and protect computer systems; techniques for the removal of malicious software; and strategies and techniques for protecting your computer and data assets.

---

Paper TT11

Monday: 2:15 - 3:05 p.m.

Room: Sapphire M

**What is the Definition of Global On-Demand Reporting Within the Pharmaceutical Industry?**

Author(s): Eric Kammer

It is not uncommon in the pharmaceutical industry to have standardized reporting for data management cleaning activities or clinical review. However, certain steps have to be taken into consideration when programming standardized reports so they run properly for all studies. One particular troublesome area is when the global reports are run on-demand by non-programmers. As new data becomes available original test cases that were used could be different and the programs could fail. This paper will identify techniques to assist programmers on what to do when developing code so programs can run successfully when unexpected surprises with data occur running SAS® reports in an on demand environment and give the customers information to assist in understanding the results.

---

Paper TT12

Monday: 3:30 - 3:50 p.m.

Room: Sapphire M

**Let the CAT Out of the Bag: String Concatenation in SAS 9**

Author(s): Josh Horstman

Are you still using TRIM, LEFT, and vertical bar operators to concatenate strings? It's time to modernize and streamline that clumsy code by using the string concatenation functions introduced in SAS 9. This paper is an overview of the CAT, CATS, CATT, and CATX functions introduced in SAS 9.0, and the new CATQ function added in version 9.2. In addition to making your code more compact and readable, this family of functions also offers some new tricks for accomplishing previously cumbersome tasks.

---

Paper TT13

Monday: 1:15 - 1:35 p.m.

Room: Sapphire M

**Internal Consistency and the Repeat-TFL Paradigm: When, Why and How to Generate Repeat Tables/Figures/Listings from Single Programs**

Author(s): Tracy Sherman, Brian Fairfield-Carter

The concept of 'repeat' tables/figures/listings (TFLs), those being groups of 'similar' output that are held to require less time and effort to produce than 'unique' TFLs, is probably familiar to most programmers. But even with the interest programmers have in improving efficiency and reducing maintenance overhead, it's surprising how often the relationship between repeat TFLs isn't exploited in code-writing. Perhaps this stems from inadequate planning (i.e. allocating work to a rapidly-assembled programming team without first assessing relatedness among TFLs), possibly combined with a lack of understanding of either the importance of grouping related output, or of the programming techniques that can be applied in producing repeat TFLs, but in any event we often see projects that follow a '1 table/1 program' (1:1) paradigm, and which treat each TFL in isolation. This paper illustrates the often-

overlooked dangers inherent to the '1:1' paradigm, and proposes as an alternative that multiple repeat TFLs should be generated by single programs, not only as a means of improving efficiency, but also to safeguard quality (particularly with regard to consistency in style and in computational methods). Simple and practical tips for deciding when groups of TFLs should be treated as repeats are offered, along with an illustration of how to set up a single program to generate multiple repeats without adding significantly to program complexity.

---

Paper TT14

Monday: 4:00 - 4:50 p.m.

Room: Sapphire M

**The Three I's of SAS® Log Messages, IMPORTANT, INTERESTING, and IRRELEVANT**

Author(s): William E Benjamin Jr

I like to think that SAS® error messages come in three flavors, IMPORTANT, INTERESTING, and IRRELEVANT. SAS calls its messages NOTES, WARNINGS, and ERRORS. I intend to show you that not all NOTES are IRRELEVANT nor are all ERRORS IMPORTANT. This paper will walk through many different scenarios and explain in detail the meaning and impact of messages presented in the SAS log. I will show you how to locate, classify, analyze, and resolve many different SAS message types. And for those brave enough I will go on to teach you how to both generate and suppress messages sent to the SAS log. This paper presents an overview of messages that can often be found in a SAS Log window or output file. The intent of this presentation is to familiarize you with common messages, the meaning of the messages, and how to determine the best way to react to the messages. Notice I said "react", not necessarily correct. Code examples and log output will be presented to aid in the explanations.

---

Paper TT15

Wednesday: 9:30 - 9:50 a.m.

Room: Sapphire L

**"Ma, How Long Do I Cook The Turkey For?"**

Author(s): David Franklin

In November it will be Thanksgiving and the cry will go out from the kitchen in some households "Ma, How Long Do I Cook The Turkey For?" This paper aims to give a light-hearted introduction to ODS RTF output using information of cooking times for that traditional of meats' at Thanksgiving, the turkey. Along the way there will be some turkey facts added in for little bit of light relief.

---

# 2015 Annual Conference of Pharmaceutical Industry SAS® Users Group



RENAISSANCE ORLANDO AT SEAWORLD

■  
Practical education for the real world

■  
Focus on pharmaceutical programming

■  
Learn about CDISC issues and applications

■  
Hear the latest on FDA guidances

■  
Network with other SAS users

■  
Provide a forum for all level of  
programmers and statisticians

**Elizabeth Dennis**  
Operations Chair

**Syamala Schoemperlen**  
Academic Chair

**May 17 - 20, 2015**

**Pharma**  
**SUG**  
**ORLANDO 2015**

The logo for Pharma SUG Orlando 2015 features the text 'Pharma SUG' in blue and green, with 'ORLANDO 2015' in blue below it. The text is overlaid on a stylized green turtle with a globe for its shell.