

Berichte aus der Statistik

**Rainer Muche,  
Rolf-Hasso Bödeker (Hrsg.)**

**KSFE 2007**

Proceedings der 11. Konferenz der SAS®-Anwender  
in Forschung und Entwicklung (KSFE)



Shaker Verlag  
Aachen 2007

**Bibliografische Information der Deutschen Nationalbibliothek**

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Copyright Shaker Verlag 2007

Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder vollständigen Wiedergabe, der Speicherung in Datenverarbeitungsanlagen und der Übersetzung, vorbehalten.

Printed in Germany.

ISBN 978-3-8322-6680-6

ISSN 1619-0963

Shaker Verlag GmbH • Postfach 101818 • 52018 Aachen

Telefon: 02407 / 95 96 - 0 • Telefax: 02407 / 95 96 - 9

Internet: [www.shaker.de](http://www.shaker.de) • E-Mail: [info@shaker.de](mailto:info@shaker.de)

## **Vorwort**

Die 11. Konferenz der SAS-Anwender in Forschung und Entwicklung (KSFE) fand am 1. und 2. März 2007 an der Universität Ulm statt. Die Universität Ulm ist eine Hochschule mit Schwerpunkten in Medizin und Naturwissenschaften. Unter dem Stichwort „Lebenswissenschaften“ werden die Kenntnisse aus den Ingenieurwissenschaften, Naturwissenschaften, Informatik und Mathematik verbunden mit der klinischen Forschung, um so die Weiterentwicklung medizinischer Diagnostik und Therapie besonders zu fördern. Dazu gehören natürlich die entsprechenden statistischen Methoden, die in der Forschung eingesetzt werden. Diese Kompetenz wird an der Universität Ulm u.a. von den Instituten für Biometrie und Epidemiologie vorgehalten und in Forschung und Lehre eingesetzt.

Die Lehre im Bereich Medizinstatistik und deren Umsetzung mit Statistiksoftware wird an der Universität Ulm in der Medizinischen Fakultät im Wesentlichen vom Institut für Biometrie durchgeführt. Dabei werden u.a. SAS-Analyst- und SAS-Syntaxkurse mehrmals jährlich angeboten. Auch in der Beratung der forschenden Kliniker und Medizinstudenten kommt eine Empfehlung für diese Software oft in Frage. Die MitarbeiterInnen der Abteilung arbeiten mit SAS als Standardwerkzeug zur Auswertung klinischer Studien. Dies gilt ebenfalls für die KollegInnen der Abteilung Epidemiologie. Daneben gibt es mit den beiden Ausbildungsstellen für Medizinische DokumentarInnen an der Schule für Med. Dokumentation und der Hochschule Ulm zwei Einrichtungen in Ulm, an denen SAS in großem Umfang gelehrt und angewandt wird.

Die 11. KSFE wurde vom Verein „Kooperation der SAS-Anwender in Forschung und Entwicklung e.V.“ und dem Institut für Biometrie der Universität Ulm gemeinsam ausgerichtet. Als Tagungsort hatte die lokale Konferenzorganisation die Medizinische Klinik auf dem Oberen Eselsberg gewählt, das einen angemessenen Rahmen in universitärem Ambiente – verbunden mit professioneller Medien- und Präsentationstechnik – für diese Konferenz bildete.

Die in der Geschichte der KSFE größte Teilnehmerzahl war ein Indiz dafür, dass das Programmkomitee unter der Leitung von Frau Dr. Carina Ortseifen eine aktuelle Konferenzthematik vorgegeben und die „richtige“ Auswahl unter den eingereichten Vortragsganmeldungen getroffen hatte: Über 400 TeilnehmerInnen von wissenschaftlichen und kommerziellen Organisationen unterschiedlichster Fachrichtungen aus dem gesamten Bundesgebiet und dem deutschsprachigen Ausland waren nach Ulm gekommen. Die große Teilnehmerzahl ist aber auch ein Beleg für die hohe Aktualität dieses anwender- und praxisorientierten Konferenztyps im Kontext der SAS-Software.

Der Vizepräsident der Universität Ulm, Herr Professor Dr. Ulrich Stadtmüller, zuständig für Lehre und Lehrstuhlinhaber für Wahrscheinlichkeitstheorie hieß die Teilnehmer herzlich willkommen. Herr Dr. Rolf-Hasso Boedeker, der Vorsitzende des Vorstandes des KSFE e.V. begrüßte ebenfalls die angereisten Teilnehmer von Seiten des Vereins und wünschte der Konferenz einen erfolgreichen Verlauf. Der Geschäftsführer von SAS Deutschland, Herr Jost Dörken gab dem Auditorium anschließend einen informativen Ausblick über zukünftige Produkte und die strategische Ausrichtung von SAS. Der Einführungsvortrag wurde von Herrn Professor Dr. Wilhelm Gaus gehalten, der eine klare Einführung in den Begriff der Wahrscheinlichkeit im medizinischen Umfeld präsentierte.

Der thematische Schwerpunkt der Konferenz „Statistik und Datenanalyse mit SAS“ spiegelt sich in den über 30 Fachvorträgen und 16 Posterpräsentationen, die auf drei Veranstaltungsreihen verteilt waren, eindrucksvoll wider. Im Vorfeld der Konferenz fanden diesjährig wieder drei Tutorien zeitlich parallel statt, die mit insgesamt 170 Teilnehmern ausgesprochen gut angenommen wurden. Begleitend fand in den Pausenbereichen eine Ausstellung mit Industriepartnern und Consultingfirmen aus dem SAS-Umfeld statt, an deren Ausstellungsständen sich während der Vortragspausen rege Informations- und Diskussionsaktivitäten entwickelten.

Den auch in diesem Jahr vergebenen Best Paper Award gewann für den besten Vortrag Dr. Rainer Kaluscha mit dem Vortrag „Datenmanagement mit Oracle, SAS, Perl und Unix-Utilities: Werkzeuge für alle Fälle“ Auf die ebenfalls prämierten Plätze 2 und 3 kamen Himmel, Michelmann, Reincke mit dem Vortrag „Vorbereitung einer halbautomatisierten Beantwortung von Gesundheitsfragen an ein medizinisches Internet-Expertenforum“ und Frau Meurer mit „Validierung von SAS-Programmen für die Auswertung und Dokumentation klinischer Daten“. Daneben wurde auf der 11. KSFE-Tagung zum ersten mal eine geführte Posterpräsentation durchgeführt mit 16 ausgewählten Postern. Das beste Poster wurde ebenfalls prämiert. Den erstmalig vergebenen Poster Award erhielt das Poster: Kögler, Ortseifen „Einführung in ein Statistikprogramm – Vergleich der SAS/Enterprise Guide 4.1 Software mit SPSS für Windows“.

An dieser Stelle sei noch einmal dem örtlichen Organisationskomitee und dessen zahlreichen Helfern, insbesondere den MitarbeiterInnen aus dem Institut für Biometrie und dem Kurs MD37 der Schule für Medizinische Dokumentation, die mit ihrem großen Einsatz wesentlich zum Erfolg dieser Konferenz beigetragen haben, ganz besonders gedankt. Auch Frau Henriette Höhle vom Universitätsrechenzentrum Heidelberg sei für ihre Mühe und Sorgfalt bei der Erstellung dieses Konferenzbandes gedankt. Zum Schluss soll noch einmal ausdrücklich den vielen Sponsoren gedankt werden, die durch ihr Engagement zum Gelingen der 11. KSFE in Ulm beigetragen haben.

Es soll aber auch nicht versäumt werden, auf die 12. Konferenz für SAS-Anwender in Forschung und Entwicklung (KSFE) am 28. und 29. Februar 2008 in Aachen hinzuweisen. Diese Konferenz wird vom dem Institut für Medizinische Statistik und dem Institut

für Statistik und Wirtschaftsmathematik der RWTH Aachen gemeinsam mit dem KSFE e.V. ausgerichtet.

Rainer Muche  
Institut für Biometrie  
Universität Ulm  
Tagungsleiter 11. KSFE

Rolf-Hasso Bödeker  
AG Med. Statistik, Inst. Med. Informatik  
Justus-Liebig-Universität Gießen  
Vorsitzender KSFE e.V.

Ulm / Gießen, im März 2007

## Inhalt

Eröffnungsvortrag: <i>W. Gaus</i>	1
<b>Wahrscheinlichkeitsaussagen in der Medizin</b>	
<i>H.-P. Altenburg, C. Ortseifen, T. Petrowitsch, G. Pfister, W. Schollenberger</i>	11
<b>Tipps und Tricks für den leichteren Umgang mit der SAS-Software</b>	
<i>H. J. Baurecht</i>	37
<b>Assoziationsstudien bei Familiendaten in der Genetischen Statistik</b>	
<i>S. Beimel</i>	41
<b>Programmierrichtlinien</b>	
<i>R. Bender</i>	55
<b>Simulation von Überlebenszeiten mit Hilfe von SAS®</b>	
<i>G. Büchele</i>	63
<b>Übereinstimmungsanalyse zweier stetiger Merkmale umgesetzt in einem SAS-Macro</b>	
<i>G. Büscher, K. Unnebrink, M. Kron</i>	71
<b>Analyse longitudinaler Daten in PROC MIXED und PROC GENMOD</b>	
<i>K. Glaß</i>	85
<b>ODS RTF – Erweiterte Möglichkeiten durch direkte RTF Befehle</b>	
<i>Th. G. Grobe</i>	97
<b>SAS-Makro-Sammlung zum Geocoding von Photos – graphische Benutzeroberfläche inklusive</b>	
<i>W. Himmel, U. Reincke, H. W. Michelmann</i>	113
<b>Semiautomatisierte Beantwortung von Gesundheitsfragen an ein medizinisches Internet-Expertenforum durch Text Mining</b>	
<i>B. Hörmann, R. Muche</i>	123
<b>Variablenselektion im linearen Regressionsmodell mit der experimentellen Prozedur PROC GLMSELECT</b>	
<i>K. Hohl</i>	137
<b>Simulationsergebnisse zum Vergleich von Ersetzungsmethoden fehlender Werte von kategorialen Variablen in SAS mit PROC MI</b>	

<i>B. P. Jäger, E. Klassen, K.-E. Biebler, P. E. Rudolph</i> <b>Der Liebermeister-Test mit SAS®</b>	<b>151</b>
<i>R. Kaluscha</i> <b>Datenmanagement mit Oracle, SAS, Perl und Unix-Utilities: Werkzeuge für alle Fälle</b>	<b>179</b>
<i>N. Kleekamp, T. Tschiesche</i> <b>Analyse des Nutzerverhaltens zum Videoserverprojekt der Friedrich-Schiller- Universität Jena</b>	<b>187</b>
<i>A. Kleiner, G. Büchele, G. Weinmayr, S. Weiland</i> <b>SAS-Makro zur Analyse von disproportionalen Substichproben am Beispiel der ISAAC Phase Two Studie</b>	<b>191</b>
<i>J. Klenk</i> <b>Überlebenszeitanalyse mit zeitabhängigen Einflussvariablen</b>	<b>207</b>
<i>C. Kögler, C. Ortseifen</i> <b>Einführung in ein Statistikprogramm – Vergleich der SAS/Enterprise Guide 4.1 Software mit SPSS für Windows</b>	<b>211</b>
<i>F. Mader, J. Saborowski</i> <b>Räumliche Analyse von Linientranssektstichproben mit Hilfe von Distance, ArcGIS und SAS</b>	<b>217</b>
<i>J. Margolis, A. Margolis</i> <b>Neue SAS-Prozedur GLMSELECT: Gehaltsanalyse und Studiengebührenmodell</b>	<b>235</b>
<i>A. Marr</i> <b>Verwendung von Proc Surveyselect im Rahmen des Matching-Verfahrens</b>	<b>249</b>
<i>N. Martinussen, L. Samsøe Schmidt, A. Harbo Poulsen, H. Lind Gregersen J. Schüz</i> <b>SAS als Basis einer Registrierungsdatenbank: Die Alternative zu Access</b>	<b>257</b>
<i>C. Meurer</i> <b>Validierung von SAS-Programmen für die Auswertung und Dokumentation klinischer Daten - Prozess, Umfang und Dokumentation der Validierung</b>	<b>263</b>

<i>R. Muche, B. Danner</i>	273
<b>Bedingte logistische Regression mit PROC LOGISTIC: Möglichkeiten seit SAS V9 und Vergleich zu PROC PHREG</b>	
<i>V. Nedbal</i>	281
<b>Integration of SAS® analytics and JMP® for life science applications: JMP® Genomics</b>	
<i>N. Nonnemacher, D. Weiland</i>	285
<b>Datenimport per SAS-Makro in papier-basierten EDC-Studien</b>	
<i>C. Ortseifen, W. F. Lesener</i>	293
<b>Konvertierung von SPSS-Format Data Files in SAS Tabellen und umgekehrt</b>	
<i>D. Rey</i>	319
<b>Automatic Best of Fit Estimation of Dose Response Curves</b>	
<i>O. Schoffer</i>	329
<b>SAS im Forschungsdatenzentrum der Statistischen Landesämter</b>	
<i>W. Schollenberger</i>	337
<b>Die Analyse des Nichts oder: Zur Bedeutung des Zufalls bei empirischer Variablenselektion</b>	
<i>F. Tanriverdio, A. Ring</i>	349
<b>Planung und Implementierung der Auswertung einer replikativen Bioäquivalenzstudie mit Hilfe der „Scaled Average Bioequivalence“</b>	
<i>J. Wisniewski, R. Meinert, K. Bhalla-Müller, D. Wachtlin</i>	365
<b>MAKS: Eine Bibliothek von SAS Makros zur Auswertung von Klinischen Studien in der CDISC SDTM Struktur</b>	
<i>M. Wodny</i>	369
<b>Parameterschätzung für Poisson-Verteilungen bei unvollständiger Beobachtung</b>	
<i>X. Yu, L. Li, S. Luke, N. Gretz</i>	381
<b>Von Daten zur Information - ein System für Mikroarray-Analyse und Datamining</b>	
<i>A. Zenk, V. Michel</i>	395
<b>Einsatz von PIAFStat und SAS zur Berichterstellung im Feldversuchswesen</b>	