EFMI
European Federation for Medical Informatics

Working Group on Medical Image Processing

TTeC 06
TROMSØ TELEMEDICINE AND eHEALTH CONFERENCE

# Workshop on Validation of Medical Image Processing Systems

**at Tromsø Telemedicine Conference TTeC 2006**
**Monday • 12 June 2006 • 09.00 – 11.30**

## Program and Handout

WELCOME
*Horsch A, Chair EFMI WG MIP*

**PRESENTATIONS**

9:00. Validation is Mandatory – Efforts of a European Working Group
*Horsch A, Punys V, Wittenberg T, EFMI*

9:15. Computer-Aided Diagnosis in Medical Imaging: From Pattern Recognition to Clinical Validation
*Wismüller A, University Munich, Germany*

9:30. Validation of 3D Segmentation Results
*Schmitt F*, Sturm P*, Priese L, University Koblenz, Germany*

9:45. On Difficulties Caused by Missing Reference Image Datasets
*Castro Martinez A, University of A Coruña, Spain*

10:00. Relevance of Validation for a Vendor of Medical Image Processing Tools
*Athelogou M, Definiens AG (presenter: Horsch)*

10:15. On Large Scale Validation of Medical Image Processing Methods
*Punys V, Co-chair EFMI WG MIP*

**DISCUSSION / PANEL**

10:30. The discussion is a main part of this workshop. It will be based on the outcome of the CARS Workshop on Validation, held in Berlin in June 2005. Main goal of the discussion is to cross-check and complete the existing concepts for validation platforms and by this to prepare the editing of a document that later can be used in discussions with industry, academia and public authorities. Ultimate goal of these efforts is to launch an international network of excellence on validation and an integrated project to set up a basic infrastructure.

**TELE-PRESENTATION (NIH, USA)**

11:00. Reference Image Database Resource (RIDER): Benchmarking Software Performance for Change Analysis Tools (CT and PET CT Lung Cancer)
*Clarke L, Croft B, Jaffe C, Sullivan D, Cancer Imaging Program, NCI, USA*

**11:30 LUNCH**

# Workshop Announcement

The Working Group Medical Image Processing (WG MIP, www.efmi-wg-mip.net) of the European Federation for Medical Informatics (EFMI, www.efmi.org) organizes a *Workshop on Validation of Medical Image Processing Systems* on June 12, 2006, in the framework of the Tromsø Telemedicine Conference (TTeC, see www.telemed.no/ttec2006).

The issue of validation platforms for image processing tasks has become additional importance in the course of research and development towards *molecular imaging* and *individualized medicine*. For example, imaging as a biomarker might soon create new methods in change measurements during drug therapies (NIST and NIH will organize a related workshop in Maryland, September this year). Together with the well-known demand for a better and more efficient culture of evaluation and validation of medical image analysis methods and systems in academia, and faster, cheaper and tool-supported approval procedures in industry, the step-by-step creation of validation platforms based on a *Public Private Partnership* approach, is of great relevance to the whole medical field. Major aspects in these efforts will rely on telematics (validation platform as part of an international research infrastructure, e.g. in the framework of multi-center studies, validated/approved image processing systems such as computer aided diagnosis tools as added value components of *eHealth and telemedicine services*), here the loop to TTeC closes.

As an expert in CAD you are kindly invited to give a presentation on validation aspects of this field at this workshop. It would be great if you could accept. Unfortunately, the working group doesn't have financial resources to cover your expenses. But maybe you can combine the journey with the EuroPACS conference taking place in Trondheim immediately after TTeC.

After the workshop it is planned to have work sessions (Monday and Tuesday) in a smaller group, in order to work up the results from the workshop and to discuss possible future strategy and collaboration.

**Background**
This is a workshop organized by the Working Group on Medical Image Processing (WG MIP) of the European Federation for Medical Informatics (EFMI) (www.efmi-wg-mip.net) on the issue of reference image databases and validation platforms for medical image processing methods and systems. The background of the workshop is the various activities that have taken place during the last years, especially in the US by the NIH, in Japan, and increasingly also in Europe, to improve the situation for validation in the field of medical image processing. The vision is to have reference image datasets with Gold standard for the respective analytical tasks (such as detection of suspicious areas in an image, segmentation and measurement of structures) as well as standardized validation software tools in order to create validation results that are comparable on an international scale. EFMI WG MIP is working on this issue since 2001, and its work has seen growing interest from academia and industry. For molecular imaging and approval scenarios in various industries image reference databases and validation tools will become a key issue. In telemedicine and eHealth settings, often images and - in the future more and more - image processing components, such as tools for computer aided diagnosis, will be involved as substantial parts of services.

**Aim**
The aim of the workshop is to shape a concrete strategy for the next steps in building such a validation platform, including probably an application for funding by the EC as network of excellence and/or an integrated project, on basis of a Puplic Private Partnership (PPP) approach.

**Concept/Participation**
Members of the EFMI Working Group on Medical Image Processing, delegates from the European Commission, and colleagues from Tromsø University Hospital UNN and the University of Tromsø UIT will be invited to participate. The workshop is open for conference participants interested in the issue. The workshop will be held in English.

**Registration**
All workshop participants must register for the TTeC at least for the Monday (day of the workshop). Registration can be done via the conference website www.telemed.no/ttec2006

# Abstracts and Speaker

## Validation is Mandatory – Efforts of a European Working Group

Horsch A, Punys V, Wittenberg T, EFMI

**Abstract.** The Working Group on Medical Image Processing (www.efmi-wg-mip.net) of the European Federation for Medical Informatics (EFMI) has started in 2002 an initiative with the aim to trigger the establishment of a reference image database for medical image processing research and development groups in order to support good validation and comparability of methods and systems. There is close contact with other initiatives, especially with the National Institute of Health and the Insight Software Consortium in the United States, as well as to industry in order to make the concept practicable. The concept of the EFMI reference image database initiative consists of the following main points:

- Create an overall, economically sustainable framework for life cycles of reference image datasets and corresponding tools meeting the demands for validation and quality control of both academia and industry in research and approval processes.
- Establish a board of experts and let them define criteria for how to assess the relevance of a medical problem with respect to the importance of image processing.
- Using the defined criteria, perform an assessment of medical problems and identify the most relevant ones with a high potential of improvement of diagnostic and treatment outcomes through the application of digital image processing methods.
- For the highly relevant problems specify the image datasets needed for scientifically sound validation and evaluation, including quality criteria and standardized data structures for annotations (Gold standards).
- Following these specifications, collect image data from image providers (single institution or a group of institutions) and prepare validated image datasets to serve as common references for research and development groups in academia and industry.
- Set up a platform for the dissemination of the reference image datasets, including bilateral co-operation agreements or contracts between provider and user with or without licensing, depending on the type of the datasets and the type of usage.
- Follow up the impact of the dissemination in terms of outcome indicators such as number and quality of published results, or number, costs and time for approval processes using the datasets, compared with the situation prior to their introduction.

During the last years conceptual and promotional work has been done by the EFMI WG MIP. Although there exists, to a certain extent, awareness of the usefulness or even necessity of the initiative, the concrete commitment to contribute is still rather limited. Academic institutions do not have the resources to set up such a framework, and therefore they have to focus on the outcomes of their own research projects, including own image data acquisition and management. Industry concentrates on the procedures required by the regulatory authorities and struggle with the threatening of their economical benefits by too long development and approval cycles. Both sides would benefit from an approved and powerful common platform for validation. Since setting up such a platform needs joint efforts from the public and the industry, the working group tries to form a strong alliance of academia and industry to create a model for a sustainable platform which shall be implemented in a common public-private effort. At the time being, an ongoing discussion among representatives of public initiatives and industry is fostered by workshops and meetings at various events (recently in USA and Europe). The aim is to coordinate and strengthen the activities and make the results available to the global community.

**Speaker.** Prof. Alexander Horsch is head of the working group for telemedicine and image processing at the Institute for Medical Statistics and Epidemiology (IMSE) of the Munich University of Technology (TUM), associate professor (Faculty of Medicine), and adjunct professor (Faculty of Science) at the University of Tromsø (UIT) for medical informatics, eHealth and telemedicine. He got his Dr.rer.nat. in computer science in 1989, and Dr.med.habil. in medical informatics from the medical faculty of the Munich University of Technology in 1999. He was head of the medical computing center of the university hospital Klinikum rechts der Isar from 1987 to 1995. Since 1992 he is lecturer in medical informatics, and he was manager of several projects in telemedicine and computer-aided diagnosis with grants from the German Ministry of Research and Technology, the Bavarian State Government, the European Union, and the German Telekom. Since September 1998 he is chair of the working group on medical image processing of the German Society of Medical Informatics,

Biometry and Epidemiology (GMDS), and since 2000 GMDS representative in the Council of the European Federation for Medical Informatics (EFMI). In 2001 he initiated and since then is chairing the EFMI Working Group on Medical Image Processing (WG MIP). Since November 2005 he is delegate of the TUM university hospital in the Telemedicine working group of the Bavarian Ministry for Research and Education. He has published more than hundred scientific papers, abstracts and chapters in journals, conference proceedings and books. He is member of the scientific program committees of Medical Informatics Europe (MIE) and Computer Assisted Radiology and Surgery (CARS), as well as for other national and international conferences, and he is reviewer for various national and international journals, conferences and scientific societies.

Dr. Vytenis Punys from Kaunas University, Lithuania, and Dr. Thomas Wittenberg from Fraunhofer Institute for Integrated Circuits in Erlangen, Germany, are co-chairs of the EFMI WG MIP.

## Computer-Aided Diagnosis in Medical Imaging: From Pattern Recognition to Clinical Validation

Wismüller A, University Munich, Germany

**Abstract.** Technical innovations in medical cross-sectional imaging have opened up new vistas for the exploration of the human body, enabling both high spatial and temporal resolution. However, these techniques have led to vast amounts of image data whose precise and reliable visual analysis by medical doctors and bioscientists requires a considerable amount of human intervention and expertise, thus resulting in a cost factor of substantial economic relevance. Hence, the computer-assisted analysis of medical image data has moved into the focus of interest as an issue of high priority research efforts. The talk covers new methods for pattern recognition and computer-aided diagnosis in the field of MRI, such as functional MRI for human brain mapping, therapy control by automatic lesion detection in multiple sclerosis, and new approaches to breast cancer diagnosis in MRI mammography. An outlook to topical projects in bioinformatics confirms the broad applicability of the presented methods. In the light of such innovative techniques for pattern recognition in biomedicine, the increasing demand for publicly accessible validation platforms is emphasized.

**Speaker.** Axel Wismueller studied medicine at the Technical University of Munich and the University of Regensburg, Germany, with study exchange programs in Switzerland and the USA (Yale University). He received his M.D. degree from the Technical University of Munich for a scientific thesis in neurology in 1992. He successfully passed the U.S. medical examinations ECFMG and FLEX. In parallel to his clinical work in internal medicine, he studied physics at the University of Munich where he received a German masters degree in theoretical physics from the University of Munich in 1996 for a scientific thesis on pattern recognition. Since 1997, he has been working as a fellow of radiology in the Department of Clinical Radiology at the University of Munich, where he founded the Digital Image Processing Group. His main research interest is focussed on innovative strategies for computer-aided diagnosis and pattern recognition in MRI data, such as functional MRI for human brain mapping or the diagnosis of breast cancer in MRI mammography. Dr. Wismueller is author of more than 80 scientific journal and conference publications related to pattern recognition in biomedicine.

## Validation of 3D Segmentation Results

Schmitt F, Sturm P, Priese L, University Koblenz, Germany

**Abstract.** To judge the quality of an image segmentation one can either compare the results of the segmentation with a known optimal segmentation, a so called ground truth, or visually evaluate the plausibility of the results. We will introduce measures to rate the quality of a segmentation if a ground truth is available and present an algorithm to automatically generate images with known ground truth suitable for evaluation of general 3D segmentation methods. In the second part of the talk, we will present our approach to validation of MR brain image segmentation. Here we used a phantom image provided by the Brainweb project and a catalog for manual judgment of segmentation quality developed together with the German Armed Forces Central Hospital (GAFCH), Koblenz, Germany.

**Speaker.** Frank Schmitt received his diploma in computer science in 2004, since 2005 he is a member of the Image Recognition Lab at the Institute for Computational Visualistics, University Koblenz-Landau. The lab is lead by Professor Lutz Priese and currently focuses its research on analysis of (mainly medical) 3D images, motion estimation in color image sequences, eye tracking and automatic extraction of striking features from color images. Our main tool in image analysis is the CSC, a powerful image segmentation method developed in our lab. Professor Priese also leads a research group working in the field of theoretical computer sciences. Here the focus of research is true-concurrency semantics of Petri nets and molecular algorithms, especially splicing systems.

## On Difficulties Caused by Missing Reference Image Datasets

Castro Martinez A, University of A Coruña, Spain

**Abstract.** One of the main problems in the development of systems for the analysis of medical images or segmentation algorithms is the set of images that is used for both the development and the testing of the system. In most cases, this images set is obtained from the historical archives of a collaborating hospital, which implies that the researchers rarely dispose of an ample set of images with which they can evaluate the used algorithms. In some cases, there is an over-adjustment of the obtained results. The first part of this lecture presents our research team and its activities. The second part briefly comments on our most important research work and the difficulties and problems that we encountered because we did not dispose of standard image reference.

**Speaker.** Alfonso Castro Martínez was born in A Coruña, Spain, in 1971. He received the University Degree in Computer Science from the University of A Coruña in 1994. Since 1995, he has been a member of the Group of Biomedical Engineering, at the Interdisciplinary Research Laboratory in Artificial Intelligence, led by Dr. Bernardino Arcay Varela. He received his Doctoral Degree for the thesis “Study and validation of advanced segmentation algorithms in medical images. Integration into an information system”, whose purpose was to fit and test the quality of various segmentation algorithms in a heterogeneous set of images. The results would be integrated into a help module for the medical expert who analyses the images in a web PACS. Dr. Castro Martínez is currently working as an Assistant Professor in the Department of Information and Communications Technologies of the Informatics Faculty. His main research area concerns segmentation algorithms for medical image analysis and their validation.

## Relevance of Validation for a Vendor of Medical Image Processing Tools

Athelogou M, Schmidt G, Definiens AG (presenter: Horsch)

**Abstract.** Image acquisition devices enable medical scientists and biologists to acquire thousands of high- and ultra high-content images every day. Information and knowledge extraction from these images has become exceedingly important. In radiology, pathology, diagnostics, toxicology, drug discovery and development processes, assay screening facilities images are used as a knowledge medium in now days than ever before. Definiens has developed the Cognition Network Technology (CNT), a general framework to develop applications to automate the analysis of such complex biological systems. It enables the user to map and program human cognitive processes using CNL (Cognition Network Language), a high-level semantic computer language in CNT that allows efficient programming to develop complex image analysis solutions. Interactive graphical representations of CNL scripting, a rich variety of classification and segmentation algorithms, variables and control structures are the main components of CNL. It is important to note, that the image analysis per se has to be considered as a key stone in the whole data analysis workflow. To validate the image analysis in the context of an automated and integrated workflow, it is essential to set up and maintain a database of annotated images. This database should comprise manually or automatically segmented and classified images whereas the data formats should be compatible to the automatically generated image analysis results to ensure automatic benchmarking. This is a need for the definition of a new standard representation of segmented and classified images. We propose to utilize for these requirements the conceptual CNT approach, whereas analyzed images are represented by a semantic, hierarchical network of image objects. The automated validation of image analysis tools using such a validated database will have a tremendous impact on

the acceptance and leverage of image analysis software and the automation of the whole data analysis workflow.

**Authors.** Dr. Maria Athelogou and Dr. Günter Schmidt are Senior Scientists at Definiens AG.

## On Large Scale Validation of Medical Image Processing Methods

Punys V, Co-chair EFMI WG MIP

**Abstract.** There is a variety of image related parameters that affect image processing methods and their results. It is desirable that medical image processing methods either have known dependencies or (the best) are independent from the acquisition parameters (e.g. spatial dependencies, scale, acquisition mode, etc.). Examples of such parameters will be given and discussed, as well as possible improvements of the Reference Image Database Concept.

**Speaker.** Vytenis Punys is a senior researcher at the Image Processing and Analysis Research Lab., Kaunas University of Technology, Lithuania. He graduated with distinction from the Moscow Institute of Electronic Engineering in 1991; and received a doctoral degree in informatics from the Institute of Mathematics & Informatics, Vilnius, in 2002. He is an expert of the Lithuanian Sate Foundation for Science and Studies since 2003. His research interests include healthcare information systems, their standards, medical image analysis. He is a co-chair of the WG MIP of the EFMI.

## TELE-PRESENTATION (NIH, USA)

# Reference Image Database Resource (RIDER): Benchmarking Software Performance for Change Analysis Tools (CT and PET CT Lung Cancer)

Clarke L, Croft B, Jaffe C, Sullivan D, Cancer Imaging Program, NCI, USA

**Abstract.** The Reference Image Database to Evaluate Response (RIDER) to therapy in lung cancer began as a highly leveraged and collaborative *pilot* project, initiated in September 2004, by the NCI's Cancer Imaging Program, NCI's Center for Bioinformatics, the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Cancer Prevention and Research Foundation, and with information technology support from the Radiological Society of North America (RSNA). The specific aims and proposed methods to develop this public resource include: (a) Development and implementation of an NCI caBIG public resource with the following specific aims; (b) Development of a broad consensus for the design and implementation of this resource by the RIDER steering committee, comprising academic researchers, program staff at NCI, members of caBIG, NIBIB, FDA (CDRH, CDER) and NIST, and implementation of the this resource using the following approaches. Building upon the success of this initial pilot phase, it is proposed that RIDER efforts be expanded from a *pilot* project to a *demonstration* project. As a demonstration project, the RIDER database would be expanded beyond a pilot project to include additional image and related metadata from modalities such as X-ray CT and extended to PET-CT as applied to lung cancer. This data will be collected from a wide range of therapy trials supported by NCI and by the pharmaceutical industry. One important goal of this effort would be to engage industry partners in the data collection, database design, and implementation to explore if the database could be useful in accelerating FDA approval and CMS reimbursement of therapeutic decisions made using software tools. – It is further proposed that this demonstration phase of the RIDER project include both continued support from the existing partners, as well as additional support for the private sector through a Public Private Partnership (PPP). Because the RIDER project is not focused on a specific clinical treatment protocol, or a specific drug, there is little or no intellectual property associated with the generation of this public resource. This demonstration project may thus provide an especially good opportunity to develop a successful PPP of interest to and supported by a broad range of stakeholders in this field, including the imaging, information technologies (IT), software and pharmaceutical industries. The PPP would be coordinated by the Foundation for NIH, a non-profit organization chartered by Congress to raise funds and establish public-private partnerships that complement and enhance NIH priorities and activities. – A longer-term goal, which underlies the use of the term "demonstration project" for this next phase of RIDER, is to use the *organizational structure* of the RIDER PPP as a model for additional projects that would employ, and further evaluate, the most highly rated software tools, using data collected from future NCI clinical trials and, more broadly, trials conducted in other NIH institutes and centers. NIBIB is especially interested an effort to engage not only NCI but other NIH ICs where imaging is being used as a clinical measure for the pathology of disease or its progression and treatment effectiveness.

**Speaker.** Dr. Clarke as of January 1999 is the Branch Chief for Imaging Technology Development for the Biomedical Imaging Program (BIP), Division of Cancer Treatment and Diagnosis, NCI, NIH. In this capacity he is responsible for development of initiatives for supporting new and emerging imaging technology, involving both academia and industry, as applied to cancer. His responsibilities also include the development of initiatives that support research resources for assessing new imaging methods including the development of international resources for evaluation of image processing algorithms. Dr. Clarke has a detail assignment at NIBIB since 2005 and a Visiting Scientists Position at NIST as of Aug 2006 and is being tasked to develop standards for biomedical imaging for therapy response from a hardware and software perspective. Before joining NCI, Dr. Clarke was a Professor of Radiology and Adjunct Professor Physics and Computer Science at the University of South Florida (USF), and Program Leader for Digital Medical Imaging Program at the H. Lee Moffitt Cancer and Research Center at USF. He has previously worked at other cancers centers at the University of Miami and the Memorial Sloan Kettering Cancer Center NYC. Dr. Clarke has been active over the last 30 years in the area of image processing for early cancer detection, cancer diagnosis and treatment response for a range of imaging modalities. He is a Fellow of the ISMRM (1994) and AAPM (1990). He graduated with a PhD in medical physics at the National University of Ireland (1978) and an MS degree in Pure and Applied Physics from Queens University of Belfast, Ireland (1968).