

ICH



2017

2017 IEEE International Conference on **Healthcare Informatics**

23-26 August 2017 • Park City, Utah, USA



Edited by Christophe Giraud-Carrier, Julio Facelli, Hiroshi Nakajima, Mollie Cummins, Gerrit Meixner
Supported by the USA National Science Foundation



[CONFERENCE INFORMATION](#)

[PAPERS BY SESSION](#)

[PAPERS BY AUTHOR](#)

[GETTING STARTED](#)

[TRADEMARKS](#)

[SEARCH](#)

Proceedings

2017 IEEE International Conference on Healthcare Informatics



Proceedings

2017 IEEE International Conference on Healthcare Informatics

23–26 August 2017
Park City, Utah



Los Alamitos, California
Washington • Tokyo



All rights reserved.

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries may photocopy beyond the limits of US copyright law, for private use of patrons, those articles in this volume that carry a code at the bottom of the first page, provided that the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

Other copying, reprint, or republication requests should be addressed to: IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, P.O. Box 133, Piscataway, NJ 08855-1331.

The papers in this book comprise the proceedings of the meeting mentioned on the cover and title page. They reflect the authors' opinions and, in the interests of timely dissemination, are published as presented and without change. Their inclusion in this publication does not necessarily constitute endorsement by the editors, the IEEE Computer Society, or the Institute of Electrical and Electronics Engineers, Inc.

IEEE Computer Society Order Number E6232
BMS Part Number CFP1744U-ART
ISBN 978-1-5090-4881-6

Additional copies may be ordered from:

IEEE Computer Society
Customer Service Center
10662 Los Vaqueros Circle
P.O. Box 3014
Los Alamitos, CA 90720-1314
Tel: + 1 800 272 6657
Fax: + 1 714 821 4641
<http://computer.org/cspress>
csbooks@computer.org

IEEE Service Center
445 Hoes Lane
P.O. Box 1331
Piscataway, NJ 08855-1331
Tel: + 1 732 981 0060
Fax: + 1 732 981 9667
[http://shop.ieee.org/store/
customer-service@ieee.org](http://shop.ieee.org/store/customer-service@ieee.org)

IEEE Computer Society
Asia/Pacific Office
Watanabe Bldg., 1-4-2
Minami-Aoyama
Minato-ku, Tokyo 107-0062
JAPAN
Tel: + 81 3 3408 3118
Fax: + 81 3 3408 3553
tokyo.ofc@computer.org

Individual paper REPRINTS may be ordered at: <reprints@computer.org>

Editorial production by Lisa O'Conner
Cover art production by Mark Bartosik



**IEEE Computer Society
Conference Publishing Services (CPS)**

<http://www.computer.org/cps>

2017 IEEE International Conference on Healthcare Informatics

ICHI 2017

Table of Contents

Welcome Message from the General Co-Chairs	xiii
Organizing Committee	xiv
Program Committee	xv
Tutorials	xvii

International Workshop on Healthcare Knowledge Discovery and Management (IHKDM)

Using Unstructured Data to Identify Readmitted Patients	1
<i>Majid Rastegar-Mojarad, Jenna K. Lovely, Joshua Pankratz, Sunghwan Sohn, Donna M. Ihrke, Amit Merchea, David W. Larson, and Hongfang Liu</i>	
Extracting Drug-Drug Interactions with Word and Character-Level Recurrent Neural Networks	5
<i>Ramakanth Kavuluru, Anthony Rios, and Tung Tran</i>	
Exploring the Performance of Stacking Classifier to Predict Depression Among the Elderly	13
<i>Eun Sung Lee</i>	
On Mapping Textual Queries to a Common Data Model	21
<i>Sijia Liu, Yanshan Wang, Na Hong, Feichen Shen, Stephen Wu, William Hersh, and Hongfang Liu</i>	
Visual FHIR: An Interactive Browser to Navigate HL7 FHIR Specification	26
<i>Na Hong, Kui Wang, Lixia Yao, and Guoqian Jiang</i>	
CVRT: Cognitive Visual Recognition Tracker	31
<i>Matthew Velazquez and Yugyung Lee</i>	
Oro Vision: Deep Learning for Classifying Orofacial Diseases	39
<i>Rajaram Anantharaman, Vidya Anantharaman, and Yugyung Lee</i>	

Paper Pession 1: Predictive Machine Learning

Predictive Modeling of Therapy Decisions in Metastatic Breast Cancer with Recurrent Neural Network Encoder and Multinomial Hierarchical Regression Decoder	46
<i>Yinchong Yang, Peter A. Fasching, and Volker Tresp</i>	
Multitask LS-Svm for Predicting Bleeding and Re-operation Due to Bleeding	56
<i>Che Ngufor, Dennis H. Murphree, Sudhindra Upadhyaya, Jyotishman Pathak, and Daryl J. Kor</i>	
A Large-Scale Exploration of Factors Affecting Hand Hygiene Compliance Using Linear Predictive Models	66
<i>Michael T. Lash, Jason Slater, Philip M. Polgreen, and Alberto M. Segre</i>	
Construction of Discharge Summaries Classifier	74
<i>Shusaku Tsumoto, Tomohiro Kimura, Haruko Iwata, and Shoji Hirano</i>	
A Machine Learning Algorithm for Identifying Atopic Dermatitis in Adults from Electronic Health Records	83
<i>Erin Gustafson, Jennifer Pacheco, Firas Wehbe, Jonathan Silverberg, and William Thompson</i>	

Paper Session 2: Systems

A Process-Oriented Approach for Supporting Clinical Decisions for Infection Management	91
<i>Bernardo Cánovas-Segura, Francesca Zerbato, Barbara Oliboni, Carlo Combi, Manuel Campos, Antonio Morales, Jose M. Juarez, Roque Marin, and Francisco Palacios</i>	
Electrooculography Based iOS Controller for Individuals with Quadriplegia or Neurodegenerative Disease	101
<i>Bryce O'Bard, Alex Larson, Joshua Herrera, Dominic Nega, and Kiran George</i>	
An Integrated Patient Genomic Information Management and Analysis System for Healthcare Professionals	107
<i>Amal Alzu'bi and Leming Zhou</i>	
Assessing STAMP EMR with Electronic Medical Record Related Incident Reports: Case Study: Manufacturer and User Facility Device Experience Database	114
<i>Fieran Mason-Blakley, Ryan Habibi, Jens Weber, and Morgan Price</i>	
Personal Health Assistance for Elderly People via Smartwatch Based Motion Analysis	124
<i>Rainer Lutze and Klemens Waldhör</i>	

Paper Session 3: HMMs, Time-Series and Optimization

Spatio-Temporal Optimization of Seasonal Vaccination Using a Metapopulation Model of Influenza	134
<i>Srinivasan Venkatramanan, Jiangzhuo Chen, Sandeep Gupta, Bryan Lewis, Madhav Marathe, Henning Mortveit, and Anil Vullikanti</i>	
Medical Workflow Modeling Using Alignment-Guided State-Splitting HMM	144
<i>Sen Yang, Moliang Zhou, Shuhong Chen, Xin Dong, Omar Ahmed, Randall S. Burd, and Ivan Marsic</i>	
Pattern Discovery from Directional High-Order Drug-Drug Interaction Relations	154
<i>Xia Ning, Titus Schleyer, Li Shen, and Lang Li</i>	
Drug-Drug Interactions (DDIs) Detection from On-Line Health Forums: Bi-Submodular Optimization (BSMO)	163
<i>Yan Hu, Rui Wang, and Feng Chen</i>	
Continuous Assessment of Children’s Emotional States Using Acoustic Analysis	171
<i>Yuan Gong and Christian Poellabauer</i>	
Multivariate Hidden Markov Models for Personal Smartphone Sensor Data: Time Series Analysis	179
<i>William Stephen van der Kamp and Nathaniel David Osgood</i>	

Paper Session 4: EHR and Sensor Data Analytics

Pattern Recognition for Automated Healthcare Assessment Using Non-invasive, Ambient Sensors	189
<i>Dino Nienhold, Rolf Dornberger, and Safak Korkut</i>	
Note Highlights: Surfacing Relevant Concepts from Unstructured Notes for Health Professionals	198
<i>Vanessa Lopez, Joao H Bettencourt-Silva, Grace McCarthy, Natasha Mulligan, Fabrizio Cucci, Stephane Deparis, Marco Sbodio, Pierpaolo Tommasi, John Segrave-Daly, Conor Cullen, Ciaran Hennessy, Beth McKeon, Karie Kelly, Russell Olsen, John Dinsmore, Anne-Marie Brady, Nagesh Yadav, and Spyros Kotoulas</i>	
Efficient Bayesian Detection of Disease Onset in Truncated Medical Data	208
<i>Bob Price, Lottie Price, Dylan Cashman, and Marzieh Nabi</i>	
Granite: Diversified, Sparse Tensor Factorization for Electronic Health Record-Based Phenotyping	214
<i>Jette Henderson, Joyce C. Ho, Abel N. Kho, Joshua C. Denny, Bradley A. Malin, Jimeng Sun, and Joydeep Ghosh</i>	
Breast Cancer Risk Prediction Using Electronic Health Records	224
<i>Yirong Wu, Elizabeth S. Burnside, Jennifer Cox, Jun Fan, Ming Yuan, Jie Yin, Peggy Peissig, Alexander Cobian, David Page, and Mark Craven</i>	

Paper Session 5: Data Mining and Analytics

Provider-Consumer Anomaly Detection for Healthcare Systems	229
<i>Luiz F. M. Carvalho, Carlos H. C. Teixeira, Wagner Meira Jr., Martin Ester, Osvaldo Carvalho, and Maria Helena Brandao</i>	
Language-Based Process Phase Detection in the Trauma Resuscitation	239
<i>Yue Gu, Xinyu Li, Shuhong Chen, Hunagcan Li, Richard A. Farneth, Ivan Marsic, and Randall S. Burd</i>	
Discovering Quantitative Temporal Functional Dependencies on Clinical Data	248
<i>Carlo Combi, Matteo Mantovani, and Pietro Sala</i>	
Evaluation of Trace Alignment Quality and its Application in Medical Process Mining	258
<i>Moliang Zhou, Sen Yang, Xinyu Li, Shuyu Lv, Shuhong Chen, Ivan Marsic, Richard A. Farneth, and Randall S. Burd</i>	
A K-Means Approach to Clustering Disease Progressions	268
<i>Duc Thanh Anh Luong and Varun Chandola</i>	

Paper Session 6: Human Factors

Fatigue Detection Model for Older Adults Using Eye-Tracking Data Gathered While Watching Video: Evaluation Against Diverse Fatiguing Tasks	275
<i>Yasunori Yamada and Masatomo Kobayashi</i>	
Making Sense in the Long Run: Long-Term Health Monitoring in Real Lives	285
<i>Jochen Meyer, Elke Beck, Merlin Wasmann, and Susanne Boll</i>	
Wearable Privacy: Skeletons in The Data Closet	295
<i>Byron Lowens, Vivian Genaro Motti, and Kelly Caine</i>	
Trend Displays to Support Critical Care: A Systematic Review	305
<i>Noa Segall, Damian Borbolla, Guilherme Del Fiol, Rosalie Waller, Thomas Reese, Paige Nesbitt, and Melanie C. Wright</i>	
User-Centered Mapping of Nurses' Workarounds to Design Principles for Interactive Systems in Home Wound Care	314
<i>Dawood Al-Masslawi, Shannon Handfield, Sidney Fels, Rodger Lea, and Leanne M. Currie</i>	

Doctoral Consortium

Towards Identifying Informal Caregivers of Alzheimer's and Dementia Patients in Social Media	324
<i>Reda Al-Bahrani, Margaret K Danilovich, Wei-keng Liao, Alok Choudhary, and Ankit Agrawal</i>	
Radiology Clinical Notes Mining Using Weighted Association Rules	325
<i>Mohammad Alodadi</i>	

Feasibility of Internet of Things Technologies to Support Aging	326
<i>Yong K. Choi</i>	
An Efficient Approach of Outbreak Preparedness for Dengue	327
<i>Nirosha Sumanasinghe Dinayadura, Armin R. Mikler, and Jayantha Muthukudage</i>	
Approximate Temporal Functional Dependencies on Clinical Data	328
<i>Matteo Mantovani</i>	
Towards Reliable Wearable-User Identification	329
<i>Sudip Vhaduri and Christian Poellabauer</i>	
Detecting and Treating Mental Illness on Social Networks	330
<i>Akkapon Wongkoblaph, Miguel A. Vadillo, and Vasa Curcin</i>	
Off-Label Drug Use Detection Based on Heterogeneous Network Mining	331
<i>Mengnan Zhao</i>	

Poster Session 1

Physiology-Aware Rural Ambulance Routing	332
<i>Mohammad Hosseini, Richard B. Berlin Jr., and Lui Sha</i>	
SNaReSim: Synthetic Nanopore Read Simulator	338
<i>Philippe C. Faucon, Parithi Balachandran, and Sharon Crook</i>	
Bridging the Gap: A Reference Information Exchange Architecture for Fusion Imaging	345
<i>Max Taggart, Mark Evans, Guilherme Del Fiol, Derek Mann, and Matt Leavitt</i>	
Computable Adherence	351
<i>Simon Diemert, Jens Weber, and Morgan Price</i>	
Integration of Accountable Care Organization and Additional Hospital Data into CMS Referral Analytics System	357
<i>Boshu Ru, Qingxin Wu, Xin Wang, Lixia Yao, and Yugang Jia</i>	
Interconnected Personal Health Record Ecosystem Using IoT Cloud Platform and HL7 FHIR	362
<i>Jaeki Hong, Peter Morris, and Jonghwa Seo</i>	
MRSA Transmission in a Personal Care Home Facility: A Spatially Explicit Agent Based Modeling Approach	368
<i>Narjes Shojaati, Mustafa Andkhoie, Osagie Osemwegie, and Nathaniel D. Osgood</i>	
Estimating Disease Onset Time by Modeling Lab Result Trajectories via Bayes Networks	374
<i>Wonsuk Oh, Pranjul Yadav, Vipin Kumar, Pedro J. Caraballo, M. Regina Castro, Michael S. Steinbach, and Gyorgy J. Simon</i>	
Deep Reinforcement Learning for Dynamic Treatment Regimes on Medical Registry Data	380
<i>Ying Liu, Brent Logan, Ning Liu, Zhiyuan Xu, Jian Tang, and Yangzhi Wang</i>	

Code2Vec: Embedding and Clustering Medical Diagnosis Data	386
<i>David Kartchner, Tanner Christensen, Jeffrey Humpherys, and Sean Wade</i>	
Deep Learning Based Recognition of Meltdown in Autistic Kids	391
<i>Venkata Sindhoor Preetham Patnam, Feba Thankachan George, Kiran George, and Abhishek Verma</i>	
An Adaptive Differential Privacy Algorithm for Range Queries over Healthcare Data	397
<i>Asma Alnemari, Carol J. Romanowski, and Rajendra K. Raj</i>	
Disease Comorbidity Linkages between MEDLINE and Patient Data	403
<i>Tejaswi Rohit Anupindi and Padmini Srinivasan</i>	
Forecasting Influenza Levels Using Real-Time Social Media Streams	409
<i>Kathy Lee, Ankit Agrawal, and Alok Choudhary</i>	
MyHealthToday: Helping Patients with their Healthschedule Using a 24-Hour Clock Visualization	415
<i>Robin De Croon, Bruno De Lemos Ribeiro Pinto Cardoso, and Katrien Verbert</i>	
A Novel Steady-State Visually Evoked Potential (SSVEP) Based Brain Computer Interface Paradigm for Disabled Individuals	421
<i>Divya Geethakumari Anil, Krupal Sureshbhai Mistry, Vaibhav Palande, and Kiran George</i>	
Unpacking Happiness: Lessons from Smartphone Photography Among College Students	429
<i>Yu Chen, Gloria Mark, Sanna Ali, and Xiaojuan Ma</i>	
Approaching the Design of an Information Display to Support Critical Care	439
<i>Thomas J. Reese, Kensaku Kawamoto, Guilherme Del Fiol, Charlene Weir, Joseph Tonna, Noa Segall, Paige Nesbitt, Rosalie Waller, Damian Borbolla, Eugene Moretti, and Melanie C. Wright</i>	
Detecting Key Drivers for Long Length of Stay in Emergency Rooms	444
<i>Eran Simhon and Yugang Jia</i>	
Managing Environments for Healthcare Information Systems Using Enterprise Application Integration	448
<i>Chang L. Kim</i>	
Structured Information Displays for the Comparison of Clinical Trials	452
<i>Jiantao Bian, Prasad Unni, Damian Borbolla, Charlene Weir, Thomas Reese, Jacob Wan, and Guilherme Del Fiol</i>	
Reference Website Use Patterns of Poison Control Center Specialists	453
<i>Jia-Wen Guo, Heather Bennett, Barbara I. Crouch, and Mollie R. Cummins</i>	
The Impact of Online Social Capital on Twitter Users At-Risk for Suicide	454
<i>Carl Hanson, Katie Meek, Emma Hunt, Megan Searles, Michael Barnes, and Christophe Giraud-Carrier</i>	

Systematic Review of mHealth Interventions Involving Fitbit Activity Tracking Devices	455
<i>Atreya Mishra, Antonio Nieto, and Spyros Kitsiou</i>	

Poster Session 2

Creating a Labeled Dataset for Medical Misinformation in Health Forums	456
<i>Alexander Kinsora, Kate Barron, Qiaozhu Mei, and V.G. Vinod Vydiswaran</i>	
Medical Concept Normalization for Online User-Generated Texts	462
<i>Kathy Lee, Sadid A Hasan, Oladimeji Farri, Alok Choudhary, and Ankit Agrawal</i>	
Exploring Timelines of Confirmed Suicide Incidents Through Social Media	470
<i>Xiaolei Huang, Linzi Xing, Jed R. Brubaker, and Michael J. Paul</i>	
An Active Learning Based Prediction of Epidural Stimulation Outcome in Spinal Cord Injury Patients Using Dynamic Sample Weighting	478
<i>Mohammad Kachuee, Lisa D. Moore, Tali Homsey, Hamidreza Ghasemi Damavandi, Babak Moatamed, Anahita Hosseini, Ruyi Huang, James Leiter, Daniel Lu, and Majid Sarrafzadeh</i>	
On Consolidated Predictive Model of the Natural History of Breast Cancer Considering Primary Tumor and Primary Distant Metastases Growth	484
<i>Ella Y. Tyuryumina and Alexey A. Neznanov</i>	
Supervised Machine Learning to Predict Follow-Up Among Adjuvant Endocrine Therapy Patients	490
<i>Morgan Harrell, Mia Levy, and Daniel Fabbri</i>	
Determining Associations with Word Embedding in Heterogeneous Network for Detecting Off-Label Drug Uses	496
<i>Christopher C. Yang and Mengnan Zhao</i>	
A Flexible Parental Engaged Consent Model for the Secondary Use of Their Infant's Physiological Data in the Neonatal Intensive Care Context	502
<i>Yvonne Choi and Carolyn McGregor</i>	
Detecting Cognitive Distortions Through Machine Learning Text Analytics	508
<i>T. Simms, C. Ramstedt, M. Rich, M. Richards, T. Martinez, and C. Giraud-Carrier</i>	
Cost Reduction via Patient Targeting and Outreach: A Statistical Approach	513
<i>David Kartchner, Andy Merrill, and Jonathan Wrathall</i>	
Catching Zika Fever: Application of Crowdsourcing and Machine Learning for Tracking Health Misinformation on Twitter	518
<i>Amira Ghenai and Yelena Mejova</i>	
Heterogenous Knowledge Discovery from Medical Data Ontologies	519
<i>Gaurang Gavai, Marzieh Nabi, Danny Bobrow, and Saied Shahraz</i>	

Single Sensor Techniques for Sleep Apnea Diagnosis Using Deep Learning	524
<i>Rahul Krishnan Pathinarupothi, Dhara Prathap J, Ekanath Srihari Rangan, Gopalakrishnan E A, Vinaykumar R, and K P Soman</i>	
Signal Analysis for Voice Evaluation in Parkinson’s Disease	530
<i>Domenico Mirarchi, Patrizia Vizza, Giuseppe Tradigo, Nicola Lombardo, Gennarina Arabia, and Pierangelo Veltri</i>	
A Robot a Day Keeps the Blues Away	536
<i>Casey C. Bennett, Selma Sabanovic, Jennifer A. Piatt, Shinichi Nagata, Lori Eldridge, and Natasha Randall</i>	
Evaluation of Sensing and Processing Parameters for Human Action Recognition	541
<i>Xiao Bo, Alan Huebner, Christian Poellabauer, Megan K. O’Brien, Chaithanya Krishna Mummidisetty, and Arun Jayaraman</i>	
Diagnosis on Mild Cognitive Impairment Patients for Alzheimer Disease with Missing Data	547
<i>Fei Gao, Jing Li, Teresa Wu, Kewei Chen, Fleming Lure, and David Weidman</i>	
Automated EEG-Based Epileptic Seizure Detection Using Deep Neural Networks	552
<i>Javad Birjandtalab, Mehrdad Heydarzadeh, and Mehrdad Nourani</i>	
Predicting High-Order Directional Drug-Drug Interaction Relations	556
<i>Xia Ning, Li Shen, and Lang Li</i>	
Position Article on Integrating Data and Model to Understand Disease Interactions	562
<i>Marzieh Nabi, Adam Arvay, Matthew Klenk, Gaurang Gavai, Daniel Bobrow, and Johan DeKleer</i>	
Extracting Intrauterine Device Usage from Clinical Texts Using Natural Language Processing	568
<i>Jianlin Shi, Danielle Mowery, Mingyuan Zhang, Jessica Sanders, Wendy Chapman, and Lori Gawron</i>	
Application of Cognitive Load Theory to the Design and Evaluation of Usability Study of mHealth Applications: Opportunities and Challenges	572
<i>Rumei Yang, Wei Wei, and Mollie R. Cummins</i>	
Semi-Synthetic Trauma Resuscitation Process Data Generator	573
<i>Sen Yang, Yichen Zhou, Yifeng Guo, Richard A. Farneth, Ivan Marsic, and Burd S. Randall</i>	
Descriptive Analysis of Communication Patterns Between a Local Poison Control Center and Community Emergency Departments	574
<i>Victoria L. Tiase, Barbara Crouch, Heather Bennett, Cindy Weng, Rumei Yang, and Mollie R. Cummins</i>	
Author Index	575

Welcome to IEEE ICHI 2017!

The IEEE International Conference on Healthcare Informatics (ICHI) series is the premier community forum concerned with the application of computer science principles, information science principles, information technology, and communication technology to address problems in healthcare, public health, and everyday wellness. It serves as a venue for discussion of innovative technical and empirical approaches, highlighting end-to-end applications, systems, and technologies, even if available only in prototype form, as well as related social and ethical implications.

The Fifth IEEE International Conference on Healthcare Informatics (ICHI 2017) took place in Park City, Utah, in the heart of the Rocky Mountains, from August 23rd to August 26th, 2017.

As in the past, the conference featured a number of keynote addresses from renowned researchers and practitioners, a multi-track technical and industry program including oral presentations and poster sessions, a panel featuring relevant agencies and foundations funding research in this area, specialized workshop and tutorials, and a doctoral consortium.

The call for papers attracted 117 submissions from 25 different countries along 3 main tracks: Human Factors, Systems, and Analytics. In addition, 11 poster submissions were received (as abstracts only). All submissions were peer-reviewed for relevance, technical soundness, originality, and overall quality. 31 papers were selected for oral presentations (27% acceptance rate), and presented in 6 non-overlapping sessions. Another 40 papers were selected for poster presentations (61% acceptance rate). Finally, 7 of the 11 poster/abstract-only submissions were also accepted. To ensure lively poster sessions, rapid-fire sessions were introduced this year giving each poster presenter 1 minute/1 slide to highlight their work in a short plenary session preceding their poster session.

The scientific tracks were complemented by an industry track featuring 2 poster presentations and 4 guest speakers from very large healthcare companies; a panel comprising representatives from the National Science Foundation, the National Institutes of Health, and the Robert Wood Johnson Foundation discussed funding opportunities; and a doctoral consortium that attracted 8 doctoral students who were able to present a poster of their research and receive individual feedback.

We are most grateful to our Program Committee consisting of over 70 members from 19 countries for their tireless work in providing meaningful and constructive reviews, selecting, and organizing papers, tutorials and workshops. We also express appreciation to our Keynote and Invited Speakers for graciously accepting our invitations, and kindly sharing their experience and expertise with the ICHI 2017 community. We are indebted to our sponsors, both corporate and government, for their financial support, including travel grants to 15 students. Finally, we express our thanks to the staff at Brigham Young University's Conferences and Workshops for their invaluable logistics support throughout.

We hope you enjoy the material found in these Proceedings, and look forward to seeing you again next year!

Julio Facelli and Christophe Giraud-Carrier
General Co-Chairs

Conference Organizing Committee

General Co-Chairs

Julio Facelli, *University of Utah, USA*

Christophe Giraud-Carrier, *Brigham Young University, USA*

Program Committee Co-Chairs

Analytics Track

Hiroshi Nakajima, *Omron Corporation, Japan*

Systems Track

Mollie Cummins, *University of Utah, USA*

Human Factors

Gerrit Meixner, *Heilbronn University, Germany*

Poster and Demo Chair

Carl Hanson, *Brigham Young University, USA*

Workshop Co-Chairs

Nitin Agarwal, *University of Arkansas at Little Rock, USA*

Amit Saha, *Wake Forest School of Medicine, USA*

Tutorial Chair

Samir E. AbdelRahman, *University of Utah, USA*

Doctoral Consortium Chair

Michael Barnes, *Brigham Young University, USA*

Industry Track Co-Chairs

Prasanna Desikan, *Blueshield of California, USA*

Steven Simske, *HP Labs, USA*

Health Data Analytics Challenge Chair

Vinod Vydiswaran, *University of Michigan, USA*

Social Media Chair

Katie Meek, *Brigham Young University, USA*

Financial and Registration Chair

Ann Baxter, *Brigham Young University, USA*

Program Committees

Systems Track

Senior Program Committee

William K. Cheung, *Hong Kong Baptist University, China*

Lei Liu, *HP Labs, USA*

Giuseppe Pozzi, *Politecnico di Milano, Italy*

Jens Weber, *University of Victoria, Canada*

Michael Chary, *New York Hospital Queens, USA*

Program Committee

Daby Sow, *IBM Research*

Harry Hochheiser, *University of Pittsburgh*

Scott Braithewaite, *Brigham Young University*

Steven Demurjian, *University of Connecticut*

Xia Hu, *Arizona State University*

Takahiro Hara, *Graduate School of Information Science and Technology, Osaka University*

Yugang Jia, *Phillips Research North America*

Fusheng Wang, *Stony Brook University*

Xia Ning, *IUPUI*

Joyce Ho, *Emory University*

Kun Bai, *IBM T. J. Watson Research Center*

Donghong Han, *Northeastern University, China*

Analytics Track

Senior Program Committee

Riccardo Bellazzi, *University of Pavia, Italy*

Aron Culotta, *Illinois Institute of Technology, USA*

Huan Liu, *Arizona State University, USA*

Aurélie Névéol, *LIMSI-CNRS, France*

Sudha Ram, *University of Arizona, USA*

Neil Smalheiser, *University of Illinois at Chicago, USA*

Zhu Tinghsao, *Chinese Academy of Sciences, China*

Yanshan Wang, *Mayo Clinic, USA*

Program Committee

Annalisa Barla, *Università di Genova*

Nicolas Bertagnolli, *3M Health Information Systems*

Michael Conway, *University of Utah*

Arianna Dagliati, *University of Pavia*

Boris Delibasic, *University of Belgrade*

Mohamed Ghalwash, *Temple University*

Mike Harmer, *Intermountain Healthcare*
David Healey, *Recursion Pharmaceuticals*
Joseph Johnson, *Brigham Young University*
John Kerley-Weekes, *SelectHealth*
Robert Leaman, *NCBI/NLM/NIH*
Paea Lependu, *Lexigram*
Aaron Ormiston, *Brigham Young University*
Petra Povalej Brzan, *University of Maribor*
Jan Ramon, *Katholieke Universiteit Leuven*
Yuval Shahar, *Ben Gurion University*
Matt Smith, *Lingotek*
Cristina Soguero Ruíz, *King Juan Carlos University*
Gregor Stiglic, *University of Maribor*
Emily Thurston, *Intermountain Healthcare*
Manabu Torii, *Kaiser Permanente*
Milan Vukicevic, *University of Belgrade*

Human Factors Track

Senior Program Committee

Sally Jo Cunningham, *Waikato University, New Zealand*
Guilherme Del Fiol, *University of Utah, USA*
Sherry Emery, *NORC at the University of Chicago, USA*
Simone Kriglstein, *Vienna University of Technology, Austria*
Anthony Maeder, *University of Western Sydney, Australia*
Jochen Meyer *OFFIS, Germany*

Program Committee

Jim Ang, *University of Kent*
Jessie Chin, *University of Illinois at Urbana Champaign*
Dominic Furniss, *University College London*
Frederik Gailly, *Ghent University*
Jenine Harris, *Washington University in St. Louis*
Chad Jensen, *Brigham Young University*
Avi Parush, *Technion University*
Christian Popow, *Medical University of Vienna, Austria*
Lucia Sacchi, *University of Pavia*
Thomas Schmidt, *The Maersk Mc-Kinney Moller Institute*
Katherine Sellen, *OCAD University*
Sei-Ching Joanna Sin, *Nanyang Technological University*
Huibert Tange, *Maastricht University*
Guenter Wallner, *University of Applied Arts Vienna*

On Consolidated Predictive Model of the Natural History of Breast Cancer Considering Primary Tumor and Primary Distant Metastases Growth

Ella Y. Tyuryumina, Alexey A. Neznanov

International Laboratory for Intelligent Systems and Structural Analysis (ISSA)

National Research University Higher School of Economics, NRU HSE

Moscow, Russia

Email: eyatyuryumina@gmail.com, aneznanov@hse.ru

Abstract—We propose a new mathematical growth model of primary tumor and primary metastases which may help to improve predicting accuracy of breast cancer process using an original mathematical model referred to CoM-IV and corresponding software. The CoM-IV model and predictive software: a) detect different growth periods of primary tumor and primary metastases; b) make forecast of patient survival; c) have higher average prediction accuracy than the other tools; d) can improve forecasts on survival of BC and facilitate optimisation of diagnostic tests. The CoM-IV enables us, for the first time, to predict the whole natural history of primary tumor and primary metastases growth on each stage (pT1, pT2, pT3, pT4) considering only on primary tumor sizes. Summarising: CoM-IV a) describes correctly primary tumor and primary distant metastases growth of IV (T1-4N0-3M1) stage with (N1-3) or without regional metastases in lymph nodes (N0); b) facilitates the understanding of the appearance period and manifestation of primary metastases.

Index Terms—breast cancer; exponential growth model; mathematical model; primary tumor; primary metastases; survival

I. INTRODUCTION

Breast cancer (BC) is the most common cancer and also the leading cause of cancer mortality in women worldwide. BC accounts for about 20-25% of all cancer types in women [1].

Finding algorithms to predict the growth of tumors has piqued the interest of researchers ever since the early days of cancer research. Many studies were carried out as an attempt to obtain reliable data on the natural history of BC growth.

Mathematical modeling can play a very important role in the prognosis of BC. Various mathematical models were built to describe primary tumor (PT) growth and distant metastases (MTS) growth separately [2].

These days, the exponential, Gompertz, logistic and von Bertalanffy models are included into a group of classical mathematical models of PT growth [3]. For the breast data, the observed linear dynamics were best captured by exponential model, which is situated for description of PT growth and, also, for secondary distant MTS growth [4]–[14]. As for Gompertz and logistic models, they are used rarely in order to describe PT growth or secondary distant MTS growth [15]–[18].

The duration of the period from the first BC cell to death refers to the natural history of BC [19]. secondary distant MTS appear in various time in different organs. The interval between removal of PT and the first clinical manifestation of MTS (MTS free survival time or nonvisible period) determined by PT size, the number of affected lymph nodes and MTS growth rate [4]–[8], [10]–[14], [16]–[18], [20]–[22]. Survival (lifetime) is the period between the date of diagnosis (TNM staging system of BC) and the date of a patient death [1]. Survival among BC patients (%) indicates the percentage of people in a study or treatment group who are alive for a given period of time after diagnosis. The percentage of patients who live at least 5-, 10-, 15-, 20-, 25- and 30-years after being treated PT is defined as 5-, 10-, 15-, 20-, 25- and 30-years observed survival rate of BC patients [1], [16].

Various forms of genetic instability and increased mutation rate can lead to a development of malignancies. I.A. Rodriguez-Brenes et al. [2] provides a complete overview of the history of mathematical models of PT growth. Nowadays, as I.A. Rodriguez-Brenes concerned, a group of classical mathematical models of PT growth consists of the exponential (with or without free initial volume), Gompertz, logistic and von Bertalanffy models [3]. For the breast data, the observed linear dynamics is best captured by exponential model that is situated for description of PT growth and, also, for secondary distant MTS growth [4]–[8], [10]–[14], [20]. As for Gompertz and logistic, models are used rarely in a description of PT and secondary distant MTS growth processes [15]–[18].

S. Benzekry et al. (2014) [3] has completed the experimental studying about the growth of primary BC and lung cancer. As researchers indicate, both of the Gompertz and exponential growth model describes primary BC most voraciously.

Unfortunately, the available papers for this do not offer mathematical models which describe MTS growth, relating with TNM classification. Consequently, a new mathematical model ought to be sought to agree with TNM classification, have higher prediction score and be independent from statistical parameters.

It is important to highlight that the natural history of BC continues after removal of PT. The next stage began with

secondary distant MTS manifestation. When the MTS reach the threshold volume, patients die from process of BC [3]–[8], [10]–[14], [16], [20], [23], [24]. All BC patients get a comprehensive treatment of PT, so the *whole natural history of BC* should include the period of secondary distant MTS growth:

- 1) the **nonvisible** period of PT growth;
- 2) the **visible** period of PT growth, diagnostics and removal of PT;
- 3) the **nonvisible** period of secondary distant MTS growth;
- 4) the **visible** period of secondary distant MTS growth, diagnostics, treatment and patient’s death.

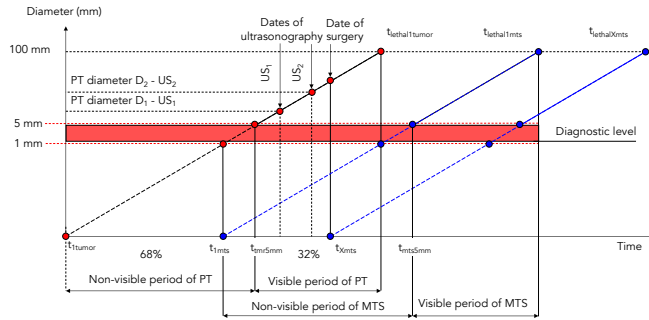


Fig. 1. Scheme of the whole natural history of BC (stage I-II) [10], [12], [13], [20], [23], [24]. As it should be highlighted, the main feature is that model describes PT growth and secondary distant MTS growth as a whole (as indivisible dependent process). Ordinate (Y): Diameter of tumor (mm). Abscissa (X): Time (years).

Legend of Fig. 1:

- t_{1tmr} — date of appearance of the first BC stem cell;
- $t_{1tmr5mm}$ — date of appearance of the visible PT with size 5 mm;
- $t_{lethal1tumor}$ — date of appearance of the lethal PT with size 100 mm (when PT reaches the threshold volume);
- t_{1mts} — date of appearance of the first MTS stem cell, which coincides with the period of 20th doubling time;
- $t_{1mts5mm}$ — date of appearance the first visible MTS of breast cancer with size 5 mm;
- $t_{lethal1mts}$ — date of appearance the first lethal MTS of BC with size 100 mm (when secondary distant MTS reaches the threshold volume);
- t_{Xmts} — date of appearance nXmts cell of BC MTS, which coincides with date of surgery;
- $t_{Xmts5mm}$ — date of appearance nXmts visible BC MTS with size 5 mm;
- $t_{lethalXmts}$ — date of appearance nXmts lethal BC MTS with size 100 mm;
- US_1 — date and sizes of the first US of PT;
- US_2 — date and sizes of the second US of PT.

Given the relation between PT and MTS, the problem of discovering BC process seems to be twofold: firstly, it is important to describe the *whole natural history of BC* to understand the process as a whole; secondly, it is necessary

to predict the period of a clinical MTS manifestation. Yet, the available papers for this do not offer mathematical growth models of MTS that relate to TNM classification. That leads to the demand of building a mathematical model that rests on exponential classical mathematical model, describes *whole natural history of BC* and corresponds to TNM classification. Moreover, the latter aspect of the problem is reflected only by statistical tools that are available as open source. In other words, a patient provides diagnostic data to predictor, and the tool calculates MTS free period and survival according to statistical data. Consequently, it is necessary to create a predictor that makes forecast for patient independently from statistical data, and requires no expensive diagnostic data. Thus, this research possesses a novelty since it is the first time the following tools for BC treatment have been proposed: a) *whole natural history of BC*; b) mathematical growth model corresponding to TNM; c) non-statistical software tool for prediction of BC developing.

To avoid terminological ambiguities, we dwell upon recalling some standard terms and TNM classification of BC (Tumor-Node-MTS) [1].

TABLE I
TNM STAGING SYSTEM

Stage	Parameter T	Parameter N	Parameter M
I	T1	N0	M0
II	T1, T2	N0, N1	M0
III	T1, T2, T3, T4	N1, N2, N3	M0
IV	any T	any N	M1

Legend of Table 1:

- parameter T — size of PT: T1 = 0.1 d ≤ 2 cm; T2 = 2 d ≤ 5 cm; T3 = d > 5 cm; T4 = spread;
- parameter N — the number of affected lymph nodes: N0: n = 0; N1: n = 1-3; N2: n = 4-9; N3: n = 10;
- parameter M — existence of distant MTS (lungs, bones, liver, etc): M0 = MTS not exist; M1 = MTS exist.

The **goal** of the research is to improve the prediction accuracy of BC process, using the original **Consolidated mathematical growth Model** of primary tumor and primary metastases (CoM-IV). To make precise the scope of the study it is necessary to fulfil several **tasks**:

1. To model the *whole natural history* of PT and MTS for stage IV;
2. To develop adequate and precise CoM-IV which reflects relations between PT and primary MTS building Consolidated mathematical growth model for *PT* and *primary MTS*, IV stage, (**CoM-IV**);
3. To analyse the CoM-IV scope of applications;
4. To implement the model as a software tool.

Practical value. As it turns out, a new software tool for prediction of BC developing can calculate more accurately: a) MTS free period; b) survival for stage IV of BC including

primary MTS and primary MTS. Moreover, the predictor can estimate a quality of treatment which was prescribed to a patient. Summarising: CoM-IV a) describes correctly PT and primary distant MTS growth of IV (T1-4N0-3M1) stage with (N1-3) or without regional metastases in lymph nodes (N0); b) facilitates the understanding of the appearance period and manifestation of primary MTS.

II. MATERIALS AND METHODS

Consolidated mathematical growth Model of PT and secondary MTS, I-II stages, (CoMPaS)

In 2015 we proposed a consolidated mathematical growth model of PT and secondary MTS (CoMPaS) that describes correctly PT growth (parameter T) as well as secondary MTS growth (parameter M), corresponds to TMN [25], [26]. Also, the CoMPaS might facilitate the survival (lifetime) and, as a consequence, make predictions of a future metastatic manifestation after removal of the PT.

It is important to define several admissions (1-3 [3]–[13], [20], [24], [27]):

- 1) the exponential growth model was used to describe "natural" growth rate of primary BC;
- 2) natural rate of secondary distant MTS is the same as "natural" growth rate of primary BC;
- 3) the period of appearance the first metastatic cell of secondary distant MTS coincides with the 20th doubling of primary BC (Fig. 2). It allows us to define the nonvisible growth period of MTS and the initial period of nonvisible MTS manifestation;
- 4) the *whole nature history* of the PT and secondary distant MTS is 60 doublings (Fig. 3).

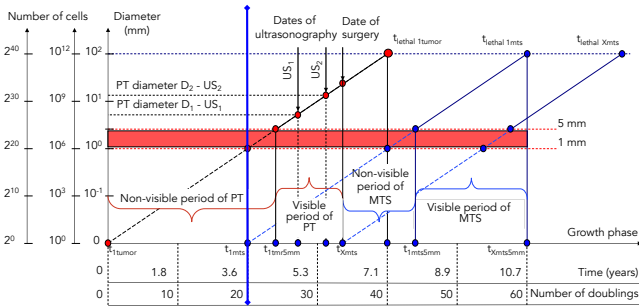


Fig. 2. The first MTS cell appears on the 20th doubling of PT

The CoMPaS is based on exponential growth model and consisted of nonlinear and linear determined equations [3]–[14], [18], [20]–[22], [24], [27]:

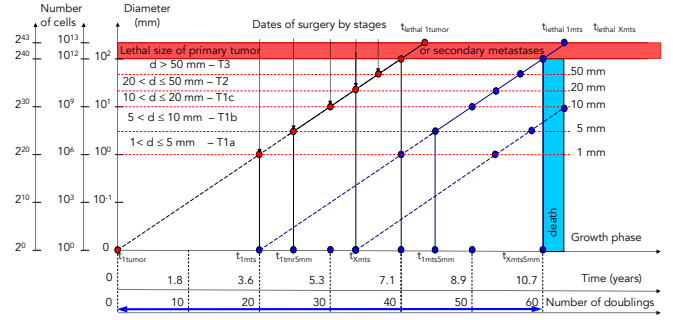


Fig. 3. The *whole nature history* of the PT and secondary distant MTS is 60 doublings

$$\left\{ \begin{array}{l} \frac{dV}{dt} = \frac{\log 2}{DT} V, \quad t \leq DT \log_2 \left(\frac{\theta DT}{\log 2} V_0 \right); \\ \frac{dV}{dt} = \theta \log V, \quad t > DT \log_2 \left(\frac{\theta DT}{\log 2} V_0 \right); \\ V(t=0) = V_0 \\ Survival = PT_{\log(V)} + Nonvis_{\log} + Vis_{\log} = 60; \\ TVDT_{non} = TVDT_{vis} = \frac{NonVis_{days} + Vis_{days}}{NonVis_{\log} + Vis_{\log}}; \end{array} \right.$$

$\frac{\log 2}{DT}$ — the fraction of proliferative cells times;
 θ — drives the linear phase;
 $PT_{\log(V)}$ — the number of PT doublings;
 $Nonvis_{\log}$ — the number of doublings for **nonvisible** growth period of MTS;
 Vis_{\log} — the number of doublings for **visible** growth period of MTS;
 $TVDT$ — tumor volume doubling time;
 60 doublings — the *whole nature growth history* of the PT and secondary distant MTS.

According to M. Schwartz (1961), the doubling time (DT) can be calculated via the measurement of tumor volume (V_1) at surgery t_1 , the first measurement of tumor volume (V_0) at diagnostic t_0 and the period between the measurements (days) ($\Delta t = t_1 - t_0$) [24]:

$$DT = \frac{\log 2 (\Delta t)}{\log V_1 - \log V_0}$$

Consolidated mathematical growth model of PT and primary MTS, IV stage, (CoM-IV)

Stage IV (T1-4N1-3M1) means that secondary distant MTS exist meanwhile PT is growing (M1 - lungs, bones, liver, etc.) [1], [28]. Five-year survival rate of stage IV is about 20-25% among BC patients [16]. Unfortunately, the available papers for this do not offer a mathematical model that describes stage IV of BC, i.e. the simultaneous growth of PT and primary distant MTS.

We propose a new mathematical growth model for PT and primary MTS. The model may help to improve predicting accuracy of BC process using an original mathematical model referred to CoM-IV and corresponding software.

The CoM-IV rests on CoMPaS and complemented by formulas that describes correcting coefficient of primary MTS growth rate (K_{pMts}) relating with PT growth rate, tumor volume doubling time of primary MTS ($TVDT_{pMts}$) relating with TVDT of PT:

$$\begin{cases} pMts_{\log(m)} = pT_{\log(V)} - 20; \\ K_{pMts} = \frac{pMts_{\log(V)}}{pMts_{\log(m)}}; \\ TVDT_{pMts} = \frac{TVDT_{pT}}{K_{pMts}}; \end{cases}$$

- $pMts_{\log(m)}$ — the number of primary MTS doublings, corresponding to IV stage;
- $pT_{\log(V)}$ — the number of PT doublings;
- K_{pMts} — correcting coefficient of primary MTS growth rate, relating with PT growth rate;
- $TVDT_{pMts}$ — tumor volume doubling time of primary MTS, relating with Tumor volume doubling time of the PT ($TVDT_{pT}$);
- $pMts_{\log(V)}$ — the number of primary MTS doublings.

III. RESULTS

Whole Natural History and calculations

The "whole natural history" of PT and **primary** MTS growth of BC for different stages (parameter T from TNM), according to CoMPaS and CoM-IV. Tab II provides calculations for the *whole natural history* of stageIV.

Legend of Tab. II:

- TNM parameters depend on PT size: T1, T2, T3, T4, N1, N2, N3, M1 [1] (see detailed description on Tab. I);
- row 1 uses data from tables of paper [28];
- row 2 is calculated from row 1;
- row 3 uses data from figure of paper [16];
- $pT_{(D)}$ means a mean size (mm) of primary tumor at surgery (removal PT);
- $pT_{\log(V)}$ means the number of doublings of PT at surgery (removal PT);
- $TVDT_{pT}$ means a mean tumor volume doubling time of PT at surgery (removal PT);
- $pMts_{(D)}$ means a mean size (mm) of primary distant MTS at surgery (removal PT);
- K_{pMts} means a mean correcting coefficient of MTS growth rate.

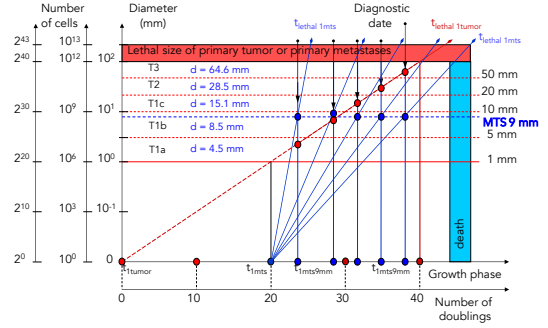


Fig. 4. Scheme of the whole natural history of BC. Stage IV

TABLE II

pT1aN1-3M1 (Fig. 5)			
$pT_{(D)}$	5.0	4.5	5.0
$pT_{\log(V)}$	26.9	26.4	26.9
$TVDT_{pT}$	80.0	80.0	80.0
$pMts_{(D)}$	5.0	9.0	50.0
$pMts_{\log(m)}$	26.9	29.4	36.9
K_{pMts}	3.9	4.6	5.34
$TVDT_{pMts}$	20.5	17.4	15.0
pT1bN1-3M1 (Fig. 6)			
$pT_{(D)}$	5.0	8.5	10.0
$pT_{\log(V)}$	26.9	29.2	29.0
$TVDT_{pT}$	75.0	75.0	75.0
$pMts_{(D)}$	5.0	9.0	50.0
$pMts_{\log(m)}$	26.9	24.9	36.9
K_{pMts}	3.9	3.2	3.72
$TVDT_{pMts}$	19.2	23.4	20.1
pT1cN1-3M1 (Fig. 7)			
$pT_{(D)}$	10.0	15.1	20.0
$pT_{\log(V)}$	29.9	31.7	32.9
$TVDT_{pT}$	70.0	70.0	70.0
$pMts_{(D)}$	5.0	9.0	50.0
$pMts_{\log(m)}$	26.9	29.4	36.9
K_{pMts}	2.72	2.52	2.86
$TVDT_{pMts}$	25.8	27.8	24.5
pT2N1-3M1 (Fig. 8)			
$pT_{(D)}$	20.0	28.5	50.0
$pT_{\log(V)}$	32.9	34.4	36.9
$TVDT_{pT}$	65.0	65.0	65.0
$pMts_{(D)}$	5.0	9.0	50.0
$pMts_{\log(m)}$	26.9	29.4	36.9
K_{pMts}	2.09	2.04	2.19
$TVDT_{pMts}$	31.2	31.8	29.7
pT3N1-3M1 (Fig. 9)			
$pT_{(D)}$	50.0	64.6	70.0
$pT_{\log(V)}$	36.9	38.0	38.3
$TVDT_{pT}$	60.0	60.0	60.0
$pMts_{(D)}$	5.0	9.0	50.0
$pMts_{\log(m)}$	26.9	29.4	36.9
K_{pMts}	1.59	1.64	2.01
$TVDT_{pMts}$	37.7	36.7	29.8

A. Limitations

Model describes **only** stage IV of breast cancer [2], [3], [14].

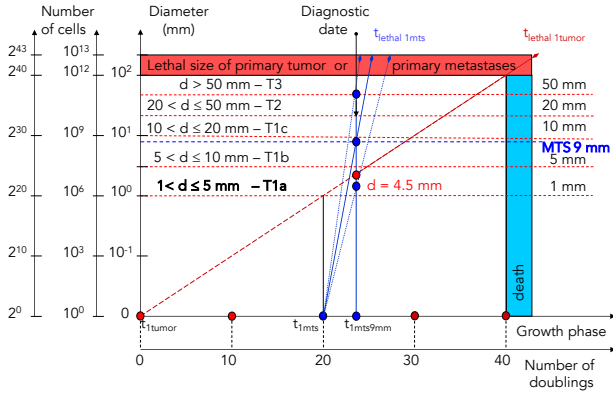


Fig. 5. pT1aN1-3M1

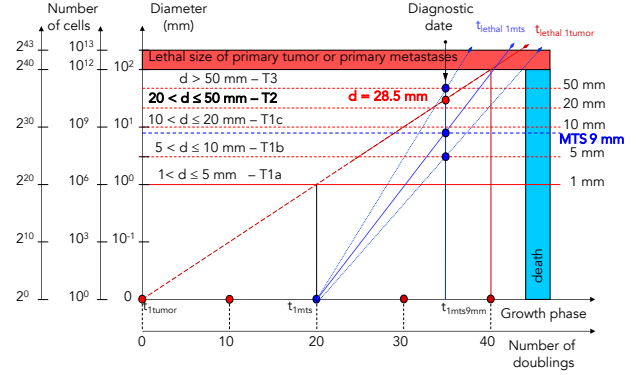


Fig. 8. pT2N1-3M1

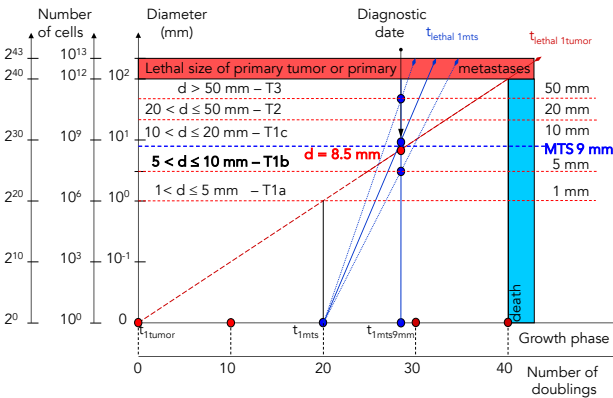


Fig. 6. pT1bN1-3M1

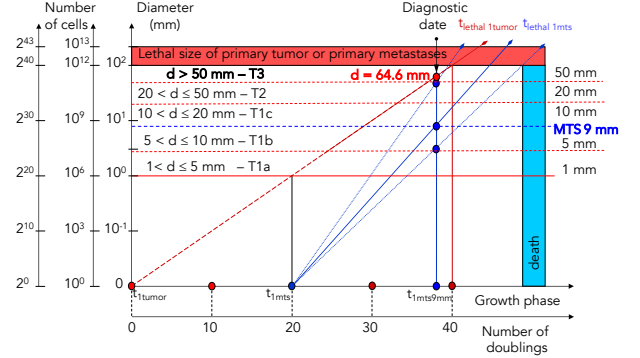


Fig. 9. pT3N1-3M1

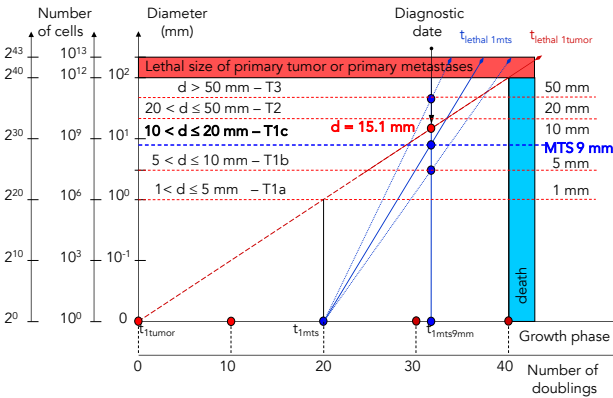


Fig. 7. pT1cN1-3M1

B. Summary

The "whole natural history" of PT and primary MTS growth of BC for different diameters of primary tumor (parameter T from TNM) was described via CoM-IV.

The CoM-IV allows us to calculate different growth periods of PT and primary MTS: 1) "non-visible period" for primary tumor; 2) "non-visible period" for primary MTS; 3) "visible period" for primary MTS.

Tab. III illustrates the variety of K_{pMts} . The higher K_{pMts} is, the shorter $TVDT_{pMts}$, the lower lifetime of patient and the worse forecast are.

TABLE III
SUMMARY

Stage	Primary MTS	K_{pMts}
pT1a (1 mm < d ≤ 5 mm)	5 mm < d ≤ 50 mm	3.90 - 5.34
pT1b (5 mm < d ≤ 10 mm)		3.72 - 3.90
pT1c (10 mm < d ≤ 20 mm)		2.72 - 2.86
pT2 (20 mm < d ≤ 50 mm)		2.09 - 2.28
pT3-4 (50 mm < d)		1.59 - 2.01

Predictor

At this stage, it is relevant to shed light on predictor specifications. The CoM-IV was implemented as a software tool. The application is build using Swift and referred as CoMPaS.

TABLE IV
INFORMATION

Developer	Ella Tyuryumina
Category	Medicine
Updated	5.04.2017
Version	2.0
Size	29 MB
Compatibility	Requires iOS 9.1 or later.
	Compatible with iPhone and iPod touch
Languages	English

INPUT DATA:

- the first ultrasound diagnostic data:
 - date (dd.mm.yyyy)
 - diameter (mm)
- the second ultrasound diagnostic data:
 - date (dd.mm.yyyy)
 - diameter (mm)
 - diameter of primary MTS (mm)

OUTPUT DATA:

- forecast:
 - the number of months
 - category of forecast:
 - * favorable
 - * mid-favorable
 - * unfavorable

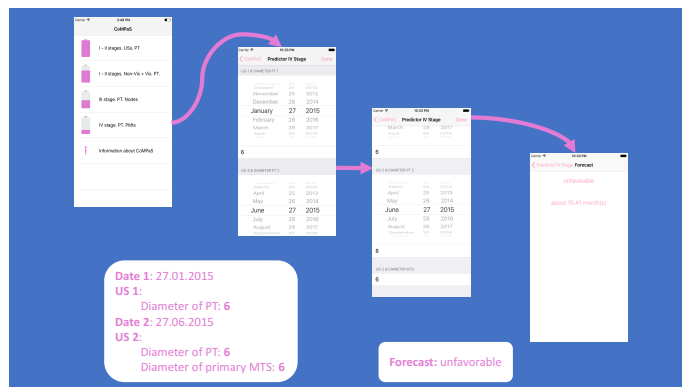


Fig. 10. Clinical example CoM-IV

It is necessary to collect predictions in one database to compare forecasts with real data and estimate effectiveness of proposed model. Consequently, the CoMPaS connects to database that allows us to test application and model.

As it turns out, the new predictive tool: 1) is a solid foundation to develop future studies of BC models; 2) does not require any expensive diagnostic tests; 3) is the first predictor which makes forecast using only current patient data, whilst the others are based on the additional statistical data.

IV. CONCLUSION

The CoM-IV: a) describes correctly PT and primary distant MTS growth of IV (T1-4N0-3M1) stage with (N1-3) or without regional MTS in lymph nodes (N0); b) facilitates the understanding of the survival period of patients with primary MTS (T1-4N1-3M1 stage IV).

The CoM-IV describes correctly (T1-4N1-3M1 stage IV): 1) the period of PT growth and corresponds to TNM classification (parameter T); 2) the period of primary distant MTS growth and 5-10 years survival of patients with stage IV of BC.

The CoM-IV calculates the variety of correcting coefficients of primary MTS growth rate, relating with PT growth rate. The constraints of application CoM-IV are imposed.

At this stage, it is relevant to dwell upon the work still to be done: 1) testing the CoM-IV on clinical data; 2) analysing forecasts statistically; 3) expanding the limits of applicability of the CoM-IV, in other words, add parameters that correspond to affected lymph nodes (N1-3); 4) implementing CoM-IV to medical practice.

REFERENCES

- [1] R. Shah, K. Rosso, and S. Nathanson, "Pathogenesis, prevention, diagnosis and treatment of breast cancer," *World J Clin Oncol*, vol. 5, no. 3, pp. 283–298, aug 2014. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pubmed/25114845>
- [2] I. Rodriguez-Brenes, N. Komarova, and D. Wodarz, "Tumor growth dynamics: insights into evolutionary processes," *Cancer*, vol. 28, no. 10, oct 2013.
- [3] S. Benzekry, C. Lamont, A. Beheshti, J. Tracz, A. and Ebos, L. Hlatky, and P. Hahnfeldt, "Classical mathematical models for description and prediction of experimental tumor growth," *PLoS Comput Biol*, vol. 10, no. 8, aug 2014.
- [4] Y. Lee and J. Spratt, "Rate of growth of soft tissue metastases of breast cancer," *Cancer*, vol. 29, no. 2, pp. 344–348, feb 1972.
- [5] A. Pearlman, "Breast cancer: Influence of growth rate on prognosis and treatment evaluation," *Cancer*, vol. 38, no. 4, pp. 1826–1833, 1976. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pubmed/991096>
- [6] P. Gullino, "Natural history of breast cancer: progression from hyperplasia to neoplasia as predicted by angiogenesis," *Cancer*, vol. 39, no. 6, pp. 2697–2703, jun 1977.
- [7] D. von Fournier, E. Weber, W. Hoeffken, M. Bauer, F. Kubli, and V. Barth.
- [8] S. Koscielny, M. Le, and M. Tubiana, "The natural history of human breast cancer: the relationship between involvement of axillary lymph nodes and the initiation of distant metastases," *Br J Cancer*, vol. 59, no. 5, pp. 775–782, may 1989. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pubmed/2736212>
- [9] S. Koscielny, M. Tubiana, and M. Le, "Breast cancer: relationship between the size of the primary tumour and the probability of metastatic dissemination," *Br J Cancer*, vol. 49, no. 6, pp. 709–715, jun 1984. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pubmed/6733019>
- [10] S. Koscielny, M. Tubiana, and A.-J. Valleron, "A simulation model of the natural history of human breast cancer," *Br J Cancer*, vol. 52, no. 4, pp. 515–524, oct 1985. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pubmed/4063132>
- [11] V. Moiseenko, "Natural history of breast cancer," *Practical oncology*, vol. 1, pp. 6–14, 2002. [Online]. Available: <http://www.practical-oncology.ru/arth009/02.pdf>
- [12] S. Friberg, "On the growth rates of human malignant tumors: implications for medical decision making," *J Oncol*, vol. 55, no. 1, pp. 1–22, aug 2005.
- [13] R. Lin and S. Plevritis, "Comparing the benefits of screening for breast cancer and lung cancer using a novel natural history model," *Cancer Causes Control*, vol. 23, no. 1, pp. 175–185, jan 2012.
- [14] R. Molina-Pena and M. Alvarez, "A simple mathematical model based on the cancer stem cell hypothesis suggests kinetic commonalities in solid tumor growth," *PLoS One*, vol. 7, no. 2, pp. 175–185, feb 2012.
- [15] L. Norton, "A gompertzian model of human breast cancer growth," *Cancer Res*, vol. 48, no. 24, pp. 7067–7071, 1988.
- [16] D. Holzel, R. Eckel, R. Emeny, and J. Engel, "Distant metastases do not metastasise," *Cancer Metastasis Rev*, vol. 29, no. 4, pp. 737–750, sep 2010.
- [17] H. Weedon-Fekjaer, B. Lindqvist, L. Vatten, O. Aalen, and S. Tretli, "Breast cancer tumor growth estimated through mammography screening data," *Breast Cancer Rev*, vol. 10, no. 3, pp. 1–13, 2008.
- [18] F. Coumans, S. Siesling, and L. Terstappen, "Detection of cancer before distant metastasis," *BMC Cancer*, vol. 13, no. 283, pp. 1–12, jun 2013.
- [19] H. Bloom, W. Richardson, and E. Harries, "Natural history of untreated breast cancer (1805-1933): comparison of untreated and treated cases according to histological grade of malignancy," *Br Med J*, vol. 2, no. 5299, pp. 213–221, 1962. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pubmed/13870135>
- [20] M. Tubiana and A. Courdi, "Cell proliferation kinetics in human solid tumors: relation to probability of metastatic dissemination and long-term survival," *Radiother Oncol*, vol. 15, no. 1, pp. 1–18, 1989. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pubmed/2664909>
- [21] S. Clare, F. Nakhliis, and J. Panetta, "Molecular biology of breast cancer metastasis: the use of mathematical models to determine relapse and to predict response to chemotherapy in breast cancer," *Breast Cancer Res*, vol. 2, no. 6, pp. 430–435, 2000. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC138666/>
- [22] H. Withers and S. Lee, "Modeling growth kinetics and statistical distribution of oligometastases," *Semin Radiat Oncol*, vol. 16, no. 2, pp. 111–119, 2006.
- [23] V. Collins, R. Loeffler, and H. Tivey, "Observations on growth rates of human tumors," *Am J Roentgen*, vol. 76, no. 5, pp. 988–1000, 1956. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2008056/>
- [24] M. Schwartz, "A biomathematical approach to clinical tumour growth," *Cancer*, vol. 14, pp. 1272–1294, nov 1961.
- [25] E. Turumina, "Consolidated mathematical growth model of breast cancer," in *School and Conference of the Kharkiv Institute for Information Transmission Problems Information Technologies and Systems*. <http://itas2015.iitp.ru/pdf/1570162553.pdf>: IITP RAS, 2015, pp. 38–51.
- [26] E. Turumina and A. Neznanov, "On consolidated predictive model of the natural history of breast cancer considering primary tumour and distant metastases growth," in *The Breast*. <http://sgbcc2017.elsevierdigital.com:ELSEVIER>, 2017, p. 106.
- [27] L. von Bertalanffy, "Quantitative laws in metabolism and growth," *Q Rev Biol*, vol. 32, no. 3, pp. 217–231, 1957. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pubmed/13485376>
- [28] J. Engel, R. Emeny, and D. Holzel, "Positive lymph nodes do not metastasise," *Cancer Metastasis Rev*, vol. 31, no. 1-2, pp. 235–246, 2012.