

# **Arianna Dagliati**

Full Name: Arianna Dagliati

Born in: Genova (GE)

Date of Birth: 11/03/1982

Address:

Cell Phone:

E-mail: [arianna.dagliati@unipv.it](mailto:arianna.dagliati@unipv.it)

## Research Statement

I am Assistant Professor (RTDa) at the Department of Electrical, Computer and Biomedical Engineering of the University of Pavia.

My research is focused on the application of Machine Learning methods within Precision Medicine and on the development of analytical approaches for identifying temporal patterns and electronic phenotypes in longitudinal clinical data.

My background is in Bioengineering and Bioinformatics, with broad experiences in applying artificial intelligence approaches for knowledge discovery, predictive modelling, temporal data mining, process mining, and software engineering.

I have an extensive history of working in multidisciplinary teams and I collaborate with a wide range of administrative, specialist and generalist clinicians, as well as public health professionals, to deliver scientific findings and to integrate algorithms in clinical decision support systems.

I hold a Master Degree and a PhD in Bioinformatics and Biomedical Engineering, both obtained at the laboratory for Biomedical Informatics Mario Stefanelli in Pavia (Italy). In 2016, I was a visiting scholar at the Department of Biostatistics and Epidemiology at the University of Pennsylvania. In 2017 I was hired by the University of Manchester, where I was promoted Research Fellow in 2018. I am a member of the 4CE Consortium for the Clinical Characterization of COVID-19 by EHR and co-led the working groups related to temporal data and unsupervised learning.

During my postdoc, at the Health eResearch and Manchester Molecular Pathology Innovation centre (UK), I was able to combine the main threads in my research: the secondary use of data to improve health and deliver better healthcare services, and the use of multiple technologies for study the heterogenous dimensions of each person's health and diseases in the context of stratified medicine.

Since I took on the role of Assistant professor (RTDa) near Department of Electrical, Computer and Biomedical Engineering of the University of Pavia, I have participated and managed work-packages dedicated to modelling techniques of temporal data within several projects. In this context, I am collaborating and provide guidance to junior staff including 2 postdocs and several master students. Starting in September 2021, I will hold the course Advanced Methods for Biomedical Data Mining in collaboration with prof. Ameen Abu-Hanna (University of Amsterdam).

My research, and my commitment to align my work to global translational medicine research priorities, embrace key steps for providing direct impact of machine learning and artificial intelligence in clinical practice: from software tools that link clinical knowledge with machine

learning-based evidences from longitudinal data, to modelling approaches to identify critical transitions in patient's histories.

I have nearly 15 years of experience in applying data mining and machine learning methods to provide insights in healthcare. I am interested in how process mining and topological data analysis techniques can be used to define novel phenotypes, which are computationally manageable and simulate disease behaviors in space and time. I have developed a careflow mining algorithm to analyze heterogeneous longitudinal data, develop a novel approach to study patient trajectories via topological approaches and I applied various statistical approaches to study diseases' evolution.

I am interested in how patients' management processes can be improved through clinical decision systems that integrates longitudinal heterogeneous data, and implements methods that illustrate changes in diseases in the context of Learning Health Systems. I have implemented a clinical decision support system that provides clinicians with a new approach for the follow up of a diabetic population, moving towards strategies focused on the continuous follow up and prevention of the disease worsening.

Since 2007 I have worked on the integration and utilization of heterogeneous types of data in different clinical context. I have implemented an ICT infrastructure to manage biological sample from a bio-bank integrated with clinical data of more than 6,500 patients with breast cancer diagnosis; a framework to collect clinical, administrative and environmental data for type 2 diabetes management and arrhythmogenic disease research; and performed large-scale analytics to study correlations among biomarkers variations and remotely sensed air pollution data.

My experience in health informatics, the scientific evidences from my publications' track, and my current research interest, interlaces with the opportunity to achieve the provision of methods and technologies, which leverage on heterogeneous and multi-scale data assets, to direct clinical practice towards precision medicine.

## Appointments

**2021 - Present** Assistant Professor (RTDa), Department of Electrical, Computer and Biomedical Engineering. University of Pavia.

Projects' participation and roles:

- Lombardy FRRB project INTESTRAT-CAD (INTEgrated STRATification Tools in Coronary Artery Disease): Co-leader (with Prof. Riccardo Bellazzi) of the work-package focused on artificial intelligence and electronic phenotyping.

- Horizon2020 project BrainTeaser (Bringing Artificial Intelligence home for a better care of amyotrophic lateral sclerosis and multiple sclerosis): Co-leader (with Prof. Riccardo Bellazzi and Dr. Roberto Bergamschi) of the work-package for the integration of environmental data in models for the evolution of MS and progression of ALS.

- 4CE Consortium for the Clinical Characterization of COVID-19: Co-leader (with prof. Shawn Murphy and prof. Tianxi Cai) of the study groups for Post-acute Sequelae of COVID-19 and Thrombotic events analyses.

- Horizon2020 project Periscope (Pan-European Response to the Impacts of COVID-19 and future Pandemics and Epidemics): Scientific collaborator for the development of modelling approaches to identify potential best practices

**2020 –2021** Research Fellow, Department of Electrical, Computer and Biomedical Engineering. University of Pavia

Lombardy FRRB project INTESTRAT-CAD (INTEgrated STRATification Tools in Coronary Artery Disease): Project and team member.

Activities: Application of Machine Learning techniques for the detection and early identification of coronary artery disease high-risk phenotypes.

Integration of multi-omics and advance coronary imaging for risk stratification in coronary atherosclerosis.

**2019 –2020** Research Fellow, Division of Informatics, Imaging & Data Sciences, The University of Manchester, UK.

Promotion to Research Fellow (August 2019).

- 2017 -2019** Post-Doctoral Research Associate, Manchester Molecular Pathology Innovation Centre, Division of Informatics, Imaging & Data Sciences, The University of Manchester, UK
- MRC funded Manchester Molecular Pathology Innovation Centre: person in charge of the Health Informatics and Machine Learning tasks.
- Activities:
- Application of advanced analytics techniques (such as Latent Class Trajectory modelling, Topological Data Analysis) to discover temporal phenotypes in inflammatory diseases and jointly study omics and clinical data.
  - Behavioral assessment of hearing through face recognition in children, in collaboration with the Manchester Centre for Audiology and Deafness;
  - UK biobank-based project for a precision medicine approach for treatment and prevention of Alzheimer's disease using statins in collaboration with Prof. Roberta Brinton (University of Arizona)
- 2016 –2017** Post-Doctoral Research Assistant, Clinical and Scientific Institute S.Maugeri of Pavia
- Activities: Implementation of software solutions to gather and analyze data within the project Inherited arrhythmia: clinical characterization, genetic geography and experimental studies in the Calabria Region isolate.
- 2009 - 2012** Data Analyst, Institute for Advanced Study of Pavia (IUSS)
- Activities: Implementation of an i2b2 based architecture to support clinical research and integrate EHR, Biobank and Hospital Information System data supporting translational research in oncology.
- 2008 - 2010** Scholarship, University of Pavia and Clinical and Scientific Institute S.Maugeri of Pavia .
- Activities: Development of a model for multi-disciplinary rehabilitative paths for the rehabilitation of breast cancer patients via muscular fatigues measurements.
- 2007 - 2009** Junior Data Analyst, Local Health Care Agency of Pavia
- Activities: Management of the Anatomy pathology Lombardy network. Pre-processing, analysis and reporting of administrative data aimed at improving the management of healthcare policy.

## Education and Qualifications

- May 2021** Associate Professor (Fascia II) Italian National Qualification  
Bando D.D. 2175/2018 – valid from 04/05/2021 to 04/05/2030  
Settore Concorsuale 09/G2 Bioingegneria
- Since 2019** New Academics and Fellows Programme at the University of Manchester  
Accredited by the Higher Education Academy (HEA) successful completion of the programme will entitle to become a Fellow of the HEA.
- Completed courses: Enhancing Research Publications, Research Grant Applications, Career Networking and Profile Raising, Undertaking Ethical Research, Media Engagement. Programme interrupted due to relocation.
- 2013-2016** Doctoral Program of Bioengineering and Bioinformatics– University of Pavia  
Thesis: *Longitudinal Data Analytics for Clinical Decision Support in Type 2 Diabetes*. Supervisor: Dr. Lucia Sacchi.  
My PhD was funded through the MOSAIC EU project, dedicated to provide new approaches to clinicians, for the diagnosis and the follow up of the chronic population with diabetes.  
My thesis embraces the following topics: (i) design of a data model to integrate temporal multivariate data from heterogeneous sources, (ii) development of predictive models and longitudinal analytics methods to determine external factors influencing the progression of Type 2 Diabetes, and (ii) integration of the developed models into a dashboard-based clinical decision system to enhance the diabetic disease management.
- During my PhD I collaborated with Professor J.H. Holmes in the University of Pennsylvania, Department of Biostatistics and Epidemiology. I have spent two months in the Department in March/May 2015 and six months in March/September 2016 as visiting scholar.
- \* The thesis was awarded the "Marco Ramoni" award for doctoral dissertation 2017, National Bioengineering Group (GNB), Italy
- \* The MOSAIC project implementation based on my thesis won the Digital Innovation in Health Care 2016/2017, Politecnico of Milano, School of Management

- 2011-2012** Master in Complex Action - International School for Advanced Studies (SISSA, Trieste)  
Fields of Interest: Analysis of complex processes and Business Administration, Food Issue. Main Courses: Accounting, Project and Risk management, Technology Transfer, Sustainability and Leadership in complex scenarios.  
Final project (Business plan): *EATIE, a platform to share healthy and sustainable food behaviors.*
- 2005-2007** Master Degree in Biomedical Engineering University of Pavia (110/110)  
Thesis: *Design and implementation of an electronic medical record to improve intensive care units' efficacy.* Supervisor: Prof. Riccardo Bellazzi.  
Thesis done at Mario Negri Institute for Pharmacological Research.
- 2001-2004** Bachelor Degree in Biomedical Engineering University of Pavia (104/110)  
Thesis: *Application of Temporal Abstraction to predict patients' stay in Intensive Care Unit.* Supervisor: Prof. Riccardo Bellazzi.
- In 2004, after my Bachelor Degree, I have spent four months in the Medical Informatics department of the Academisch Medisch Centrum of Amsterdam to finalize a data mining project on monitoring data in intensive care units, under supervision of Dr. N. Peek and Professor A. Abu-Hanna.
- 1997 -2001** Classical High School A. Doria, Novi Ligure, Italy (100/100)  
Final project: *The role of women in history, from Calpurnia to Partisan Women.*  
Main Subjects: Latin and Greek Literature, Italian Literature, History and Philosophy, Art History.

## **Prizes from Scientific Societies**

- PhD thesis awarded with the "Marco Ramoni" award for doctoral dissertation 2017, National Bioengineering Group (GNB), Italy
- I received the 2021 AIME (society of AI in medicine) Rising Star prize.

## **Other evidence of Academic and Professional Standing**

### **Board membership**

- Appointed as AIME (society of AI in medicine) Board Member since June 2021

### **Program committee member**

- Part of the program committee of Artificial Intelligence in Medicine (since 2017)
- Part of the program committee of IEEE International Conference on Healthcare Informatics (since 2017)
- Part of the program committee of IEEE CBMS International Symposium on Computer-Based Medical Systems (since 2019)
- Review for Frontiers in Digital Health (since 2020)

### **Peer Review**

Since 2015, I have reviewed 100 manuscripts for several journals including Journal of Biomedical Informatics, BMC Medical Informatics and Decision Making, IEEE Journal of Biomedical and Health Informatics, International Journal of Medical Informatics, Plos One, Artificial Intelligence in Medicine, BMJ Open, Briefing in Bioinformatics, Scientific Report, Patterns.

Peer Review Publons profile is available: <https://publons.com/researcher/1361396/arianna-dagliati/peer-review/>

### **Grant Review**

- 2019, Innova Molise Review of 2 grant proposals– Technology Innovation EU funds
- 2019, Barts Charity Grant Review - Strategic Research Grant
- 2020, Israel Science Foundation Grant Review - Personal Research Grants



## **Leadership and Management Roles**

- I was in charge for the ONCO-i2b2 project's management, and I provided administrative support (2009-2012)
- I was part of the organization and scientific committee of the Workshop NETTAB 2011, dedicated to the Clinical Bioinformatics.
- I co-organized the course "Clinical Bioinformatics" for Medical Students near "Collegio Ghislieri" of Pavia from 2013 to 2016.
- As part of the MOSIAC EU Project's consortium, I was responsible for deliverables' production for the work package dedicated to predictive modelling and temporal analytics. I actively participate in the scientific management of the project during internal and plenary meetings.
- Since 2020 I have organized the "Machine Learning Fridays" series of seminars near the Department of Electrical, Computer and Biomedical Engineering of the University of Pavia

## **Knowledge and Technology Transfer**

- I collaborated within the Process mining for healthcare initiative to draft their manifesto, and to integrate my algorithm in a collaborative software library (pMineR)
- I have collaborated with the i2b2 team and participated to annual meetings with the academic research group for the development of innovative solutions for Translational Research in Medicine (i2b2 and Shrine)

# Teaching

## Formal Assignments

- Since the 2021/22 academic year – teach the course Advanced Methods for Biomedical Data Mining near the University of Pavia (Italy). 3 CFU
- Since 2020/2021 academic year – teach the course General Informatics for “Corsi di Laurea delle Professioni Sanitarie Tecniche” near the University of Pavia (Italy). 1 CFU.
- 2019 - teaching on Causal Reasoning, as part of the Biomedical Modelling for Health Data course (IIDS67642) of the MSc in Health Data Science (UK)
- 2019 – teaching on Process mining – UCL, London, as part of the MSc Health Informatics joint award between the University of Manchester and University College London (UK)

## Assistance

- 2021 – assistant teaching in the Advanced Methods in Biomedical Data Mining. R practical part and missing data. The course was co-held by Prof. Lucia Sacchi (University of Pavia) and Prof. John H. Holmes (Perelman School of Medicine, UPenn)
- 2016 – assistant teaching in the Advanced Methods in Biomedical Data Mining, R practical part. The course was co-held by Prof. Riccardo Bellazzi (University of Pavia) and Prof. Stefano Monti (Boston University)

# Supervision of Research Students

## PhD students

- In 2018 I have co-supervised a PhD student (Dr. Kathryn A McGurk) for 3 months, in collaboration with Prof. Anthony Whetton (University of Manchester) on the use of Machine Learning techniques to analyse proteomics data.  
*Research output:* Kathryn A McGurk, Arianna Dagliati, et al., The use of missing values in proteomic data-independent acquisition mass spectrometry to enable disease activity discrimination, *Bioinformatics*, Volume 36, Issue 7, 1 April 2020,
- In 2018 I have co-supervised a visiting PhD student (Dr. Lorenzo Chiudinelli) from University of Pavia, in collaboration with Prof. Niles Peek (University of Manchester) on the development of a software tool to derive temporal phenotypes in breast cancer with process mining techniques.  
*Research output:* Chiudinelli L, Dagliati A, et al.. Mining post-surgical care processes in breast cancer patients. *Artif Intell Med.* 2020 May;105:101855. doi: 10.1016/j.artmed.2020.101855. Epub 2020

## Master and Bachelor Students

- I am currently supervising two Master and one Bachelor Student at the University of Pavia.
- In 2020/2021 I have supervised one master student and two bachelor students at the University of Pavia, in collaboration with Prof. Riccardo Bellazzi (University of Pavia)
- In 2019/2020 I have supervised two Master Students in collaboration with Prof. Riccardo Bellazzi (University of Pavia) on the application of topological data analysis on proteomics data and on the analysis of time series for behavioural pattern recognition.  
*Research Output:* Arioli A, Dagliati A, Geary B, Peek N, Kalra PA, Whetton AD, Geifman N. OptiMissP: A dashboard to assess missingness in proteomic data-independent acquisition mass spectrometry. PLoS One. 15;16(4):e0249771. (2021)
- In 2018 I have supervised a Master Students in collaboration with Prof. Carlo Combi (University of Verona) on the application of latent class mixed models to retrieve temporal phenotypes.  
*Research output:* B Amico, A Dagliati, D Plant, A Barton, N Peek, N Geifman. A Dashboard for Latent Class Trajectory Modeling: Application in Rheumatoid Arthritis. Studies in health technology and informatics 264, 911-9151 (2019)
- During my PhD (2013/2017) I have co-supervised five bachelor students and two master students near the Laboratory for Biomedical Informatics "Mario Stefanelli", University of Pavia in collaboration with Prof. Riccardo Bellazzi and Prof. Lucia Sacchi

## Scientific collaborations and participation to national and international research projects

<b>Project Name</b>	<b>Funding Agency</b>	<b>Role in the Project</b>	<b>Participants</b>
<b>Rete anatomie patologiche lombarde (2007-2009)</b>	Regione Lombardia	Junior researcher	ASL Pavia, Policlinico San Matteo Pavia, Istituto Nazionale dei Tumori
<b>Onco-OnReahb, Definizione di un modello di percorso riabilitativo multidisciplinare e costituzione di un osservatorio nazionale di riabilitazione oncologica (2008-2010)</b>	Programma straordinario per la ricerca oncologica, Ministero della Salute	Member of the research team	Fondazione S. Maugeri - Clinica del lavoro e della riabilitazione, Universita' degli Studi di Pavia
<b>Metodologie E Tecnologie Bioinformatiche Per L'integrazione Di Informazioni Cliniche E Conoscenze Biologiche A Supporto Della Ricerca Translazionale In Oncologia (Onco-I2b2) (2009-2012)</b>	Regione Lombardia	Member of the research team	Fondazione S. Maugeri - Clinica del lavoro e della riabilitazione, I.U.S.S. - Istituto Universitario di Studi Superiori - PAVIA, Universita' degli Studi di Pavia

<b>Project Name</b>	<b>Funding Agency</b>	<b>Role in the Project</b>	<b>Participants</b>
<b>MOSAIC (Models and simulation techniques for discovering diabetes influence factors) (2013-1016)</b>	EU, FP7	Member of the research team	Medtronic Ibérica SA Università degli Studi di Padova, Universidad Politecnica de Madrid, Fondazione Salvatore Maugeri, AEDEC, Lund University, TSB, Folkhalsan Research Center, University of Athens
<b>Manchester Molecular Pathology Innovation Centre (MMPathIc) (2017-2020)</b>	MRC UK	Member of the research team	University of Manchester
<b>MATURA (MAXimising Therapeutic Utility for Rheumatoid Arthritis) (2017-2020)</b>	Versus Arthritis, MRC	Scientific Collaborator	Queen Mary University London, University of Manchester
<b>A precision medicine approach for treatment and prevention of Alzheimer's disease using statins (2018-2021)</b>	University of Arizona	Member of the research team	University of Arizona, University of Manchester
<b>NIHR Biomedical Research Centre: Hearing Health Theme (2018-2020)</b>	NIHR	Member of the research team – Referent for Machine Learning Research	University of Manchester

<b>Project Name</b>	<b>Funding Agency</b>	<b>Role in the Project</b>	<b>Participants</b>
<b>4CE Consortium for Clinical Characterization of COVID-19 by EHR data-driven studies (since 2020)</b>	NIH funds requested	Member of the consortium, Work-package co-leader	The project is coordinated by the Harvard University. Full participants list: <a href="https://www.covidclinical.net/members/index.html">https://www.covidclinical.net/members/index.html</a>
<b>INTEGRATED-CAD (INTEGRATED STRATIFICATION TOOLS IN CORONARY ARTERY DISEASE) (2019-2024)</b>	Fondazione Regionale per la Ricerca Biomedica	Work-package co-leader	Centro Cardiologico Monzino, Istituto Clinico Humanitas, IFOM, Policlinico San Matteo, University of Pavia
<b>Brainteaser, Bringing artificial intelligence home for a better care of amyotrophic lateral sclerosis and multiple sclerosis. (2020-2025)</b>	H2020	Work-package co-leader	Universidad Politecnica De Madrid , Universita Degli Studi Di Padova , Universita Degli Studi Di Torino , Instituto De Medicina Molecular Joao Lobo Antunes , Servicio Madrilenio De Salud , Fondazione Istituto Neurologico Nazionale Casimiro Mondino
<b>Periscope (Pan-European Response to the Impacts of COVID-19 and future Pandemics and Epidemics) (2020-2023)</b>	H2020	Scientific Collaborator	The Project involves 31 participants. I collaborate with Fondazione Irccs Policlinico San Matteo, Assistance Publique Hopitaux De Paris And Universidad Politecnica De Madrid

## Grants proposals

- PI for “Ricerca biomedica condotta da giovani ricercatori - 2021” of Fondazione Cariplo, project “TEMporal Phenotyping approaches to model and explain LatE post-traumatic Seizures after traumatic brain injury (TEMPLES)”. Requested fund EUR 250,000. Not successful.
- PI for Chronic Renal Insufficiency Cohort (CRIC) Opportunity Pool, project “A Topological Approach to identifying Temporal Cardiac Trans-Phenotypes in Chronic Renal Insufficiency” Requested fund USD 75,000. Not successful.
- Co-I for William Demant Oticon Foundation, project: Next-generation paediatric hearing assessment: optimising behavioural testing and development of automated classification of responses. Fund GBP 250,000. Successful.
- Co-I for Versus Arthritis, project: Targeting tumour necrosis factor inhibitors to the patients with rheumatoid arthritis most likely to benefit from them. Not successful.

## Conference and workshop presentations

### Invited Speaker

- American Diabetes Association 80th Scientific Session 2020. “Use of AI to Screen for Complications” within the Artificial Intelligence, Machine Learning, and Diabetes session.
- University of Manchester, 8 Mar 2018 Informatics for Stratified Medicine and Biomarker Discovery: Research symposium. Unsupervised sub-groups discovery using Topological Data Analysis

### Panel presentations

- AMIA 2021 Annual Symposium. “Temporal Phenotypic Pathways of Post-Acute Sequelae of SARS-CoV-2 by an International Consortium for Clinical Characterization of COVID-19 (4CE)”. Panel participants: Shawn N. Murphy (Harvard Medical School), Hossein Estiri (Harvard Medical School), Riccardo Bellazzi (University of Pavia), John H. Holmes (Perelman School of Medicine, UPenn)

### Oral presentations

- IEEE BHI, Valencia, 1- 4 Giugno 2014 Dagliati A, Sacchi L, Cerra C, Leporati P, De Cata P, Chiovato L, Holmes, JH, Bellazzi R. Temporal data mining and process mining techniques to identify cardiovascular risk-associated clinical pathways in Type 2 diabetes patients. 2014 IEEE-EMBS International Conference on Biomedical and Health Informatics, BHI 2014; Valencia; Spain; 1 - 4 June 2014; Article number 6864348, Pages 240-243.
- IEEE BHI, Valencia, 1- 4 Giugno 2014 Dagliati A, Sacchi L, Bucalo M, Segagni D, Zarkogianni K, Millana AM, Cancela J, Sambo F, Fico G, Barreira MT, Cerra C, Nikita K, Cobelli C, Chiovato L, Arredondo MT, Bellazzi R. A data gathering framework to collect Type 2 diabetes patients data. 2014 IEEE-EMBS International Conference on Biomedical and Health Informatics, BHI 2014; Valencia; Spain; 1 - 4 June 2014; 2014, Article number 6864349, Pages 244-247.
- Nov 15, 2014, Washington, DC AMIA Workshop on Data Mining for Medical Informatics 2014 (<http://www.dmmh.org/dmmi14>). Mining Careflows of Breast Cancer Patients.
- Nov 15, 2014, Washington, DC AMIA Workshop on Data Mining for Medical Informatics 2014 (<http://www.dmmh.org/dmmi14>). Risk-Associated Temporal Clinical pathways in T2D Patients.
- HEALTHINF 2016 - 9th International Conference on Health Informatics, Proceedings; Part of 9th International Joint Conference on Biomedical Engineering Systems and Technologies.



Learning T2D evolving complexity from EMR and administrative data by means of Continuous time Bayesian networks. BIOSTEC 2016 2016, Pages 338-344

- AMIA Symposium 2016, Nov 12 - 16, Chicago. Hierarchical Bayesian Logistic Regression to forecast metabolic control in type 2 DM patients. Annual Symposium proceedings. AMIA Symposium. Volume 2016, 2016, Pages 470-479
- Jodrell Bank Centre for Astrophysics (JBCA) Machine Learning Workshop, University of Manchester, 13 Nov 2018 15 Nov 2018. Use of Topological Data Analysis approaches to tackle precision medicine issues. Application of unsupervised methods to identify temporal phenotypes. (<https://jbca-machinelearning.github.io/index.html>)
- 17th Conference on Artificial Intelligence in Medicine, AIME 2019; Poznan; Poland; 26 June 2019 through 29 June 2019; Code 227519 Inferring temporal phenotypes with topological data analysis and pseudo time-series. Volume 11526 LNAI, 2019, Pages 399-409
- MEDINFO 2019. Aug 25-30 2019 Lyon. A Dashboard for Latent Class Trajectory Modeling: Application in Rheumatoid Arthritis. Volume 264: MEDINFO 2019: Health and Wellbeing e-Networks for All

## **Poster presentations**

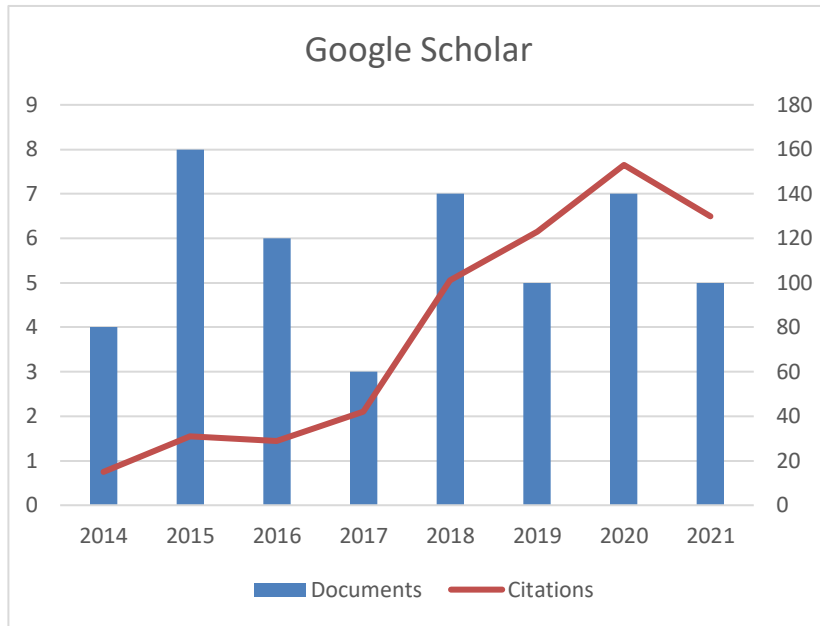
- 14th Annual Diabetes Technology Meeting, from November 6 to November 8, 2014, Bethesda. Cardiovascular-Risk-Associated Qualitative Pathways in Type 2 Diabetes Patients  
A Multivariate Data-Driven Model to Investigate the Arising of Complications in Type 2 Diabetes Patients  
Temporal Association Rules for Stratification of Type 2 Diabetes Patients
- 15th Annual Diabetes Technology Meeting, from October 22 to October 24, 2015, Bethesda. Metformin Exposure Patterns Are Related to Type 2 Diabetes Nephropathy  
Predicting Microvascular Complications from Type 2 Diabetes Retrospective Data

## **Outreach and public engagement**

- I was invited to present at the Barclays AI Frenzy x Manchester Futurists, AI in Healthcare in March 2018 as part of One HealthTech, where I discussed “Machine Learning for Digital Health Science”
- Engagement in public dissemination activities: British Science Week March 2019, Manchester Science Festival October 2019
- I participated to the UK Health Data Analytics Network workshop on health data analytics to build the UK-HDAN Roadmap

# Publications

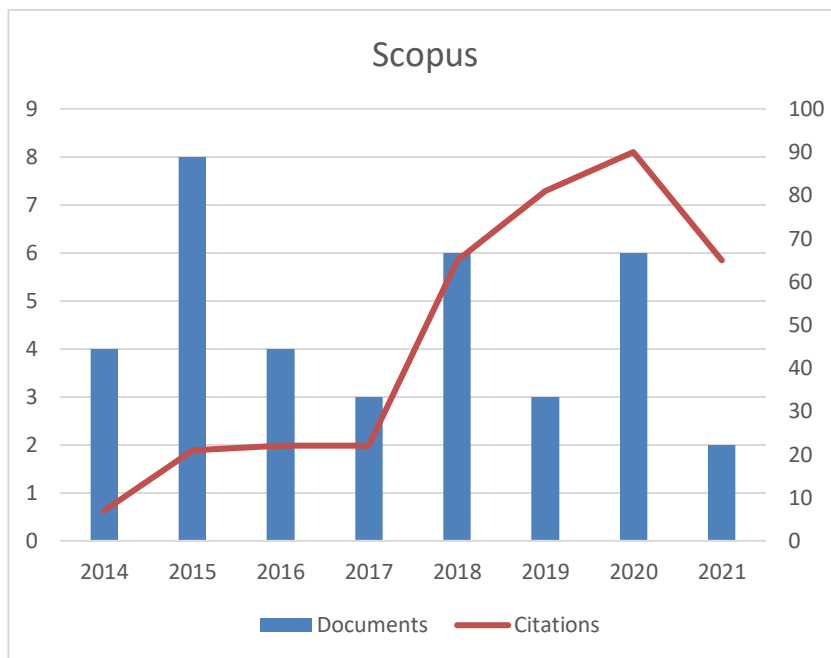
## Citation Metrics



	<i>All</i>	<i>Since 2016</i>
<i>Citations</i>	654	589
<i>h-index</i>	15	14
<i>i10-index</i>	21	19

Citation from Google Scholar (accessed on 15<sup>th</sup> July 2021)

<https://scholar.google.co.uk/citations?user=YHGB5CUAAAAJ&hl=en&oi=ao>



	<i>All</i>	<i>Since 2016</i>
<i>Citations</i>	380	348
<i>h-index</i>	13	10

Documents and Citation from Scopus (accessed on 15<sup>th</sup> July 2021)

<https://www.scopus.com/authid/detail.uri?authorId=57192824262>

## Overview of the submitted publications to be evaluated

1. **Dagliati A**, Malovini A, Tibollo V, Bellazzi R. Health informatics and EHR to support clinical research in the COVID-19 pandemic: an overview. *Brief. Bioinform.* 2021.

An overview of **health informatics** tools and international initiatives to address some of the major challenges of the **COVID-19 pandemic**. The paper provides directions regarding interoperability, processes awareness, data protection and governance.

2. Arioli A, **Dagliati A**, Geary B, Peek N, Kalra PA, Whetton AD, et al. OptiMissP: A dashboard to assess missingness in proteomic data-independent acquisition mass spectrometry. *PLoS One*. Public Library of Science; 2021.

The implementation of an interactive tool (R Shiny) to assess **missingness** patterns in **proteomic** data. The tool includes features, based on topological representation of the data, for the visualization of missingness patterns. Case study in **chronic kidney disease**.

3. **Dagliati A**, Geifman N, Peek N, Holmes JH, Sacchi L, Bellazzi R, et al. Using topological data analysis and pseudo time series to infer temporal phenotypes from electronic health records. *Artif Intell Med.* 2020.

The development of an innovative approach to **mine patients' trajectories** and stratify them accordingly. Case study in **Type 2 Diabetes**. The algorithm has been made available to the 4CE consortium to mine COVID-19 patients' trajectories.

4. **Dagliati A**, Nicora G\*, Vitali F\*, Geifman N, Bellazzi R. Integrated Multi-Omics Analyses in Oncology: A Review of Machine Learning Methods and Tools. *Front. Oncol.* 2020. \* equal first authors.

A **systematic review** of state-of-the-art **machine learning** approaches to analyze **cancer** omics data. All the first authors contributed in defining the paper organization, the research query and to the discussion. I have curated the Network-Based Methods and Clustering sections.

5. Chiudinelli L, **Dagliati A**, Tibollo V, Albasini S, Geifman N, Peek N, et al. Mining post-surgical care processes in breast cancer patients. *Artif Intell Med.* 2020.

The application of the **careflow mining** algorithm (see paper n.10) along with a **topic modeling** pre-processing of the EHR data. The paper reports some methodological improvement of the algorithm and several advancements for its validation in the context of **breast cancer**. The first author is a PhD student I co-supervised.

6. **Dagliati A**, Plant D\*, Nair N, Jani M, Amico B, Peek N, et al. Latent class trajectory modelling of 2-components-DAS28 identifies multiple rheumatoid arthritis phenotypes of response to biologic disease modifying anti-rheumatic drugs. *Arthritis Rheumatol*. 2020. \* equal first authors.

Application of a robust statistical framework based on **mixed effect models** to retrieve and validate **rheumatoid arthritis** activity **trajectories**. In the paper I have closely collaborated with the equal first author to the study design. I have defined and implemented the statistical analysis pipeline.

7. McGurk KA, **Dagliati A**, Chiasserini D, Lee D, Plant D, Baricevic-Jones I, et al. The use of missing values in proteomic data-independent acquisition mass spectrometry to enable disease activity discrimination. *Bioinformatics*. 2019.

Study of the **missingness patterns** and assessment of their impact for **proteomic** discovery within a longitudinal study of disease activity. Use Case in **rheumatoid arthritis**. The first author is a PhD student I co-supervised.

8. **Dagliati A**, Sacchi L, Tibollo V, Cogni G, Teliti M, Martinez-Millana A, et al. A dashboard-based system for supporting diabetes care. *J Am Med Informatics Assoc*. 2018.

The implementation of the **careflow mining** algorithm (see paper n.10) within a **clinical decision system** for the management of **type 2 diabetic** patients. The paper describes all the implementation phases of the system, from the design, based on learning health system conceptualization, to its validation.

9. Orphanou K, **Dagliati A**, Sacchi L, Stassopoulou A, Keravnou E, Bellazzi R. Incorporating repeating temporal association rules in Naïve Bayes classifiers for coronary heart disease diagnosis. *J Biomed Inform*. 2018.

Integration of **temporal association rules** approaches with a naïve Bayes classifier for the for diagnosis of **coronary heart disease**. The framework well-handled class imbalance issues.

10. **Dagliati A**, Sacchi L, Zambelli A, Tibollo V, Pavesi L, Holmes JH, et al. Temporal electronic phenotyping by mining careflows of breast cancer patients. *J Biomed Inform*. 2017.

Presentation of a novel careflow mining algorithm to perform **electronic phenotyping** and its initial implementation. Case study in **breast cancer**. As reported in later papers, the algorithm has been then included in a support system and exploited in several another context. It has been also included in the pMiner R package for clinical process mining.

11. **Dagliati A**, Marini S, Sacchi L, Cogni G, Marsida T, Tibollo V, et al. Machine Learning Methods to Predict Diabetes Complications. *J Diabetes Sci Technol*. 2017.

A pipeline to exploit **machine learning** approaches to build predictive models from EHR data to infer the onset of **type 2 diabetes** microvascular complications. The paper also discusses the importance of deliver explainable results to clinicals.

12. **Dagliati A**, Marinoni A, Cerra C, Decata P, Chiovato L, Gamba P, et al. Integration of Administrative, Clinical, and Environmental Data to Support the Management of Type 2 Diabetes Mellitus: From Satellites to Clinical Care. *J Diabetes Sci Technol*. 2015.

The paper is the results of a collaboration with a Telecommunications and Remote Sensing Laboratory to **integrate clinical and environmental data** for pattern recognition. Case study in **type 2 diabetes**, to correlate disease variations with seasonal pollution trends.

## Peer reviewed journal publications

1. Dagliati A, Malovini A, Tibollo V, Bellazzi R. Health informatics and EHR to support clinical research in the COVID-19 pandemic: an overview. *Brief Bioinform.* 22;22(2):812-822. (2021)
2. Arioli A, Dagliati A, Geary B, Peek N, Kalra PA, Whetton AD, Geifman N. OptiMissP: A dashboard to assess missingness in proteomic data-independent acquisition mass spectrometry. *PLoS One.* 15;16(4):e0249771. (2021)
3. Dagliati A, Peek N, Brinton RD, Geifman N. Sex and APOE genotype differences related to statin use in the aging population. *Alzheimers Dement (N Y).* May 2;7(1):e12156. (2021)
4. Bourgeois FT et al. As member of the Consortium for Clinical Characterization of COVID-19 by EHR (4CE). International Analysis of Electronic Health Records of Children and Youth Hospitalized With COVID-19 Infection in 6 Countries. *JAMA Netw Open.* Jun 1;4(6):e2112596. (2021)
5. Weber GM, Hong C, Palmer NP, Avillach P, Murphy SN, Gutiérrez-Sacristán A, Xia Z, Serret-Larmande A, Neuraz A, Omenn GS, Visweswaran S, Klann JG, South AM, Loh NHW, Cannataro M, Beaulieu-Jones BK, Bellazzi R, Agapito G, Alessiani M, Aronow BJ, Bell DS, Bellasi A, Benoit V, Beraghi M, Boeker M, Booth J, Bosari S, Bourgeois FT, Brown NW, Bucalo M, Chiovato L, Chiudinelli L, Dagliati A, et al. International Comparisons of Harmonized Laboratory Value Trajectories to Predict Severe COVID-19: Leveraging the 4CE Collaborative Across 342 Hospitals and 6 Countries: A Retrospective Cohort Study. *medRxiv [Preprint].* Feb 5:2020.12.16.20247684. Currently under review at JMIR.
6. Dagliati, A, Geifman, N., Peek, N., Holmes, J.H., Sacchi, L., Bellazzi, R., Sajjadi, S.E., Tucker, A. Using topological data analysis and pseudo time series to infer temporal phenotypes from electronic health records. *Artif Intell Med.* 108; 101930 (2020)
7. Le Sueur H, Dagliati A, Buchan I, et al. Pride and prejudice - What can we learn from peer review? *Med Teach.* 1-7. (2020)
8. Nicora G, Vitali F, Dagliati A, Geifman N, Bellazzi R. Integrated Multi-Omics Analyses in Oncology: A Review of Machine Learning Methods and Tools. *Front Oncol* ;10:1030. (2020)

9. Dagliati A, Plant D, Nair N, et al. Latent class trajectory modelling of 2-components-DAS28 identifies multiple rheumatoid arthritis phenotypes of response to biologic disease modifying anti-rheumatic drugs. *Arthritis Rheumatol* ;10.1002/art.41379. (2020)
10. Chiudinelli L, Dagliati A, Tibollo V, et al. Mining post-surgical care processes in breast cancer patients. *Artif Intell Med* ;105:101855 (2020)
11. McGurk KA, Dagliati A, Chiasserini D, et al. The use of missing values in proteomic data-independent acquisition mass spectrometry to enable disease activity discrimination. *Bioinformatics*. 36(7):2217-2223. (2020)
12. G Fico, L Hernandez, J Cancela, A Dagliati, L Sacchi, A Martinez-Millana et. Al.. What do healthcare professionals need to turn risk models for type 2 diabetes into usable computerized clinical decision support systems? Lessons learned from the MOSAIC project. *BMC medical informatics and decision making* 19 (1), 163 (2019)
13. Dagliati, A. et al. A dashboard-based system for supporting diabetes care. *Journal of the American Medical Informatics Association* 25-5, 538-547 (2018).
14. Orphanou K, Dagliati A, L Sacchi L. et al, Incorporating repeating temporal association rules in Naïve Bayes classifiers for coronary heart disease diagnosis. *Journal of biomedical informatics* 81, 74-82 (2018)
15. Dagliati, A. et al. Big Data as a Driver for Clinical Decision Support Systems: A Learning Health Systems Perspective. *Front. Digit. Humanity*. 5,8 (2018).
16. Dagliati, V. Tibollo, G. Cogni, L. Chiovato, R. Bellazzi, and L. Sacchi, “Careflow Mining Techniques to Explore Type 2 Diabetes Evolution,” *Journal of Diabetes Science and Technology*, vol. 12, no. 2, pp. 251–259, (2018).
17. Teliti, M, G Cogni, L Sacchi, A Dagliati, S Marini, V Tibollo, P De Cata, R Bellazzi, and L Chiovato., “Risk factors for the development of micro-vascular complications of type 2 diabetes in a single-centre cohort of patients.,” *Diabetes and Vascular Disease Research* 15, 5, 424-432. (2018).
18. Pinto S, K Waki, L Chiovato, P De Cata, Dagliati A, V Tibollo, G Ruvolo Smartphone-Based Self-Management of Non-Insulin-Dependent Diabetes: A Japanese System at Use by an Italian Patients’ Cohort *Journal of Diabetes Science and Technology*, 12-4, 903-904 (2018)

19. Dagliati, A. Temporal electronic phenotyping by mining careflows of breast cancer patients. *Journal of biomedical informatics*, 136–147 (2017).
20. Dagliati et al., Machine Learning Methods to Predict Diabetes Complications, *Journal of Diabetes Science and Technology* 12-2, 295-302 (2017).
21. Dagliati, A. et al. Integration of Administrative, Clinical, and Environmental Data to Support the Management of Type 2 Diabetes Mellitus: From Satellites to Clinical Care. *Journal of Diabetes Science and Technology*. 10, 19–26 (2015).
22. R. Bellazzi, A. Dagliati, L. Sacchi, and D. Segagni, “Big Data Technologies: New Opportunities for Diabetes Management,” *Journal of Diabetes Science and Technology* 9-5, 1119-1125 (2015).
23. Sacchi, L., Dagliati, A., Bellazzi, R. Analyzing complex patients’ temporal histories: New frontiers in temporal data mining. *Methods in Molecular Biology*, 1246 89-105 (2015)
24. D. Segagni, V. Tibollo, A. Dagliati, A. Zambelli, S. G. Priori, and R. Bellazzi. An ICT infrastructure to integrate clinical and molecular data in oncology research. *BMC Bioinformatics* 13, S5 (2012).
25. D. Segagni, V. Tibollo, A. Dagliati, C. Napolitano, S. G. Priori, and R. Bellazzi. CARDIO-i2b2: Integrating arrhythmogenic disease data in i2b2. *Studies in Health Technology and Informatics* 180, 1126–1128 (2012)

### **Selected peer reviewed conference publications**

26. A Dagliati, N Geifman, N Peek, JH Holmes, L Sacchi, SE Sajjadi, A Tucker. Inferring Temporal Phenotypes with Topological Data Analysis and Pseudo Time-Series. *Conference on Artificial Intelligence in Medicine in Europe*, 399-40 (2019)
27. B Amico, A Dagliati, D Plant, A Barton, N Peek, N Geifman. A Dashboard for Latent Class Trajectory Modeling: Application in Rheumatoid Arthritis. *Studies in health technology and informatics* 264, 911-915 1 (2019)
28. Dagliati, A. et al. Hierarchical Bayesian Logistic Regression to forecast metabolic control in type 2 DM patients. *AMIA Annu Symp Proc* 470–479 (2017).
29. Gatta, R., Vallati, M., Lenkowicz, J., Rojas, E., Damiani, A., Sacchi, L., De Bari, B., Dagliati, A., Fernandez-Llatas, C., Montesi, M., Marchetti, A., Castellano, M., Valentini, V. Generating and comparing knowledge graphs of medical processes using pMineR *Proceedings of the Knowledge Capture Conference, K-CAP* (2017)



30. Marini S, Dagliati A, Sacchi L, B. R. Learning T2D evolving complexity from EMR and administrative data by means of Continuous time Bayesian networks. In SciTePress. HEALTHINF 2016 - 9th International Conference on Health Informatics 338–344 (2016).
31. Marinoni, A., Dagliati, A., Bellazzi, R., Gamba, P. Inferring air quality maps from remotely sensed data to exploit geo-referenced clinical onsets: The Pavia 2013 case. International Geoscience and Remote Sensing Symposium (IGARSS), November 2015 – 3937-3940 (2015)
32. Segagni, D., Sacchi, L., Dagliati, A., Tibollo, V., Loporati, P., De Cata, P., Chiovato, L., Bellazzi, R. Improving Clinical Decisions on T2DM Patients Integrating Clinical, Administrative and Environmental Data Studies in Health Technology and Informatics, 216, 682-686. (2015)
33. Fico, G., Cancela, J., Arredondo, M.T., Dagliati, A., Sacchi, L., Segagni, D., Millana, A.M., Fernandez-Llatas, C., Traver, V., Sambo, F., Facchinetti, A., Verdu, J., Guillén, A., Bellazzi, R., Cobelli, C. User requirements for incorporating diabetes modelling techniques in disease management tools IFMBE Proceedings, 45, 992-995 (2015)
34. Dagliati, A. et al. A data gathering framework to collect Type 2 diabetes patients' data. IEEE-EMBS Int. Conf. Biomed. Heal. Informatics 244–247 (2014).
35. Dagliati, A. et al. Temporal data mining and process mining techniques to identify cardiovascular risk-associated clinical pathways in Type 2 diabetes patients. IEEE-EMBS Int. Conf. Biomed. Heal. Informatics 240–243 (2014). doi:10.1109/BHI.2014.6864348
36. Fernandez-Llatas, C., Sacchi, L., Benedi, J.M., Dagliati, A., Traver, V., Bellazzi, R. Temporal abstractions to enrich Activity-Based Process Mining corpus with clinical time series IEEE-EMBS International Conference on Biomedical and Health Informatics, BHI 2014, art. no. 6864481, 785-788 (2014)
37. Hernandez, L., Onieva, J., Fico, G., Cancela, J., Dagliati, A., Bucalo, M., Sacchi, L., Bellazzi, R., Arredondo, M.T. A proposal of architecture to share patients' data out of healthcare settings for research purposes IEEE-EMBS International Conference on Biomedical and Health Informatics, art. no. 6864482, 789-792 (2014)
38. Segagni D, Tibollo V, Dagliati A, Malovini A, Zambelli A, Napolitano C, Priori SG, Bellazzi R. Clinical and research data integration: the i2b2-FSM experience. AMIA Jt Summits Transl Sci Proc. 2013 Mar 18:239-40 (2013).
39. Bellazzi R, Dagliati A, F Licciulli, P Romano, L Varesio, M Cannataro IT solutions for integrating clinical and molecular data to support biomedical research: from biobanks to knowledge repositories ACM SIGBioinformatics Record 3 (1), 18-21 (2013)

Pavia 20/07/2021

Arianna Dagliati

*Arianna Dagliati*