GLOBAL HEALTH 2020

The Ninth International Conference on Global Health Challenges

ISBN: 978-1-61208-817-4

October 25 - 29, 2020

GLOBAL HEALTH 2020 Editors

Birgit Gersbeck-Schierholz, Leibniz Universität Hannover, Germany

Lorena Parra Boronat, Instituto Madrileño de Investigación y Desarrollo Rural,

Agrario y Alimentario and Universitat Politecnica de Valencia, Spain

GLOBAL HEALTH 2020

Forward

The Ninth International Conference on Global Health Challenges (GLOBAL HEALTH 2020), held on October 22-29, 2020, took a global perspective on population health, from national to cross-country approaches, multiplatform technologies, from drug design to medicine accessibility, everything under mobile, ubiquitous, and personalized characteristics of new age population.

Recent advances in technology and computational science influenced a large spectrum of branches in approaching population health. Despite significant progresses, many challenges exist, including health informatics, cross-country platforms interoperability, system and laws harmonization, protection of health data, practical solutions, accessibility to health services, and many others. Along with technological progress, personalized medicine, ambient assistance and pervasive health complement patient needs. A combination of classical and information-driven approach is developing now, where diagnosis systems, data protection mechanisms, remote assistance and hospital-processes are converging.

The conference had the following tracks:

- Health and Wellness Informatics
- Challenges
- Alternative
- Global

We take here the opportunity to warmly thank all the members of the GLOBAL HEALTH 2020 technical program committee, as well as all the reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors that dedicated much of their time and effort to contribute to GLOBAL HEALTH 2020. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

We also gratefully thank the members of the GLOBAL HEALTH 2020 organizing committee for their help in handling the logistics and for their work that made this professional meeting a success.

We hope that GLOBAL HEALTH 2020 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress in the domain of global health.

GLOBAL HEALTH 2020 Steering Committee

Matthieu-P. Schapranow, Hasso Plattner Institute für Digital Engineering gGmbH, Germany
Lorena Parra Boronat, Instituto Madrileño de Investigación y Desarrollo Rural, Agrario y Alimentario and Universitat Politecnica de Valencia, Spain

GLOBAL HEALTH 2020 Publicity Chair

Daniel Andoni Basterrechea, Universitat Politecnica de Valencia, Spain

GLOBAL HEALTH 2020

COMMITTEE

GLOBAL HEALTH Steering Committee

Matthieu-P. Schapranow, Hasso-Plattner-Institut für Digital Engineering gGmbH, Germany
Lorena Parra Boronat, Instituto Madrileño de Investigación y Desarrollo Rural, Agrario y Alimentario and Universitat Politecnica de Valencia, Spain

GLOBAL HEALTH 2020 Publicity Chair

Daniel Andoni Basterrechea, Universitat Politecnica de Valencia, Spain

GLOBAL HEALTH 2020 Technical Program Committee

Karin Ahlin, Mid Sweden University, Sweden
Kyota Aoki, Utsunomiya University, Japan
Marzieh Ayati, University of Texas Rio Grande Valley, USA
Samuel Bamgbose, Prairie View A&M University, USA
Panagiotis D. Bamidis, School of Medicine - Aristotle University of Thessaloniki, Greece
Oresti Banos, University of Granada, Spain
Mohamed Basel Almourad, Zayed University, Dubai, UAE
Vilmos Bilicki, University of Szeged, Hungary
Ozgu Can, Ege University, Turkey
Chih-Lin Chi, University of Minnesota, USA
James J. Cimino, Informatics Institute | Medicine University of Alabama at Birmingham - School of Medicine, USA
Javier Civit, COBER S.L., Seville, Spain
Sandra Costanzo, University of Calabria, Italy
Steven A. Demurjian, The University of Connecticut, USA
John Gialelis, University of Patras, Greece
Marie-Pierre Gleizes, IRIT, France
David Greenhalgh, University of Strathclyde, Glasgow, UK
Vida Groznik, University of Primorska & University of Ljubljana, Slovenia
Teresa Guarda, Universidad Estatal Peninsula Santa Elena - UPSE / Universidad de las Fuerzas Armadas - ESPE / ALGORITMI Research Centre | ESPE | UPSE, Ecuador
Albert Guvenis, Institute of Biomedical Engineering | Boğaziçi Üniversity, Turkey
Ying-Feng Hsu, Osaka University, Japan
Zina Ibrahim, King's College London, UK
Kamran Iqbal, University of Arkansas at Little Rock, USA
Anđela Jakšić Stojanović, Universit Donja Gorica, Montenegro
Ravi Janardan, University of Minnesota-Twin Cities, USA
Ah-Choo Koo, Digital Cities Research Institute | Multimedia University, Malaysia
Evgeniy Krastev, Sofia University St. Kliment Ohridsky, Bulgaria
Li-Hui Lee, National Taipei University of Nursing and Health Sciences, Taipei, Taiwan
Lenka Lhotska, Czech Technical University in Prague, Czech Republic

Igor Ljubi, University College for Applied Computer Engineering - Zagreb, Croatia
Matthias Loebe, Miami Transplant Institute | University of Miami Miller School of Medicine | Jackson Health Systems, USA
Fenglong Ma, Pennsylvania State University, USA
Flora Malamateniou, University of Piraeus, Greece
Stefano Mariani, Politecnico di Milano, Italy
Dulani Meedeniya, University of Moratuwa, Sri Lanka
Brendan O'Flynn, Tyndall Microsystems | University College Cork, Ireland
G M Atiqur Rahaman, Khulna University, Bangladesh
Alvaro Rocha, AISTI & University of Coimbra, Portugal
Carsten Röcker, FraunhoferIOSB-INA, Germany
Nancy Rodriguez, LIRMM | Université de Montpellier | CNRS, France
Magda Rosenmoller, IESE Business School, Barcelona, Spain
Nandan Sarkar, Thapar Institute of Engineering and Technology, India
Laura Sbaffi, University of Sheffield, UK
Matthieu-P. Schapranow, Hasso-Plattner-Institut für Digital Engineering gGmbH, Germany
Pietro Siciliano, Institute for Microelectronics and Microsystems (IMM-CNR), Lecce, Italy
Joaquim P. Silva, Polythecnic Institute of Cávado and Ave, Portugal
Marcelo Simão de Vasconcellos**,** Oswaldo Cruz Foundation (Fiocruz), Rio de Janeiro, Brazil
Patrícia Isabel Sousa Trindade Silva Leite, Polythecnic Institute of Cávado and Ave, Portugal
Paulo Sousa Teixeira, Polythecnic Institute of Cávado and Ave, Portugal
Vasile Stoicu-Tivadar, University Politehnica Timișoara, Romania
Vahid Taslimitehrani, Realtor.com, USA
Masatsugu Tsuji, Kobe International University / Osaka University, Japan
Gary Ushaw, Newcastle University, UK
Santosh Vijaykumar, Northumbria University, UK
Massimo Villari, Università di Messina, Italy
Vivian Vimarlund, Linköping University, Sweden
Dongwen Wang, Arizona State University, USA
Mudasser F. Wyne, National University, USA
Erliang Zeng, University of Iowa, USA
Qiang Zhu, University of Michigan - Dearborn, USA

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

Semantic Interoperability of Medical Information Systems and Scientific Repositories

Dimitar Tcharaktchiev
Medical University
Department of Medical Informatics, USHATE
Sofia, Bulgaria
email: dimitardt@gmail.com

Evgeniy Krastev
Faculty of Mathematics and Informatics
Sofia University St. Kliment Ohridsky
Sofia, Bulgaria
email: eck@fmi.uni-sofia.bg

Emanuil Markov
Technical University
Sofia, Bulgaria
email: emospy@gmail.com

Ivan Ivanov
Technical University
Sofia, Bulgaria
email: ivan.evgeniev@gmail.com

***Abstract*—This study presents the implementation of European health informatics standards in the Hospital Information System (HIS) and the Scientific Repositories employed in the University Specialized Hospital for Active treatment of Endocrinology. It has the purpose to demonstrate the usage of the archetype paradigm in achieving semantic interoperability in clinical data exchange. This approach enables to transfer several medico-administrative, clinical and laboratory data from the Hospital Information System to the Register of rare endocrine diseases. The Orphanet nomenclature of rare diseases as well as international Logical Observation Identifiers Names and Codes (LOINC) maintained by the Regenstrief Institute, and Systematized Nomenclature Of Medicine Clinical Terms (SNOMED-CT) codes are used for this purpose. Initial computer experiments are demonstrated and discussed.**

Keywords-Semantic interoperability; archetype object model; Health informatics standards; Medical Information Systems; Hospital Information Systems; Clinical registries.

I. INTRODUCTION

This study presents the implementation of European health informatics standards in the Hospital Information System (HIS) and the Scientific Repositories working in the University Specialized Hospital for Active Treatment of Endocrinology, (USHATE), and the transfer of clinical data between them preserving their clinical context.

II. METHODOLOGY

The patient is in the center of the integration of all clinical and administrative data. Documentation and messages conforming to the United Nations rules for Electronic Data Interchange for Administration, Commerce and Transport (UN/EDIFACT) are created [1]. The standard EN ISO 13606 for Electronic Health Record (EHR) communication and archetype paradigm is applied (see Figure 1). Roger's definition for Minimum Basic Data Set (MBDS) is adopted including core of information with the most commonly available set of items and most extensive range of usages [1]. The applied standards and nomenclatures support the semantic interoperability between the systems.

III. RESULTS

Several medico-administrative, clinical and laboratory data can be transferred from the Hospital Information System to the Register of rare endocrine diseases, keeping the context of their registration, structuring the measured results, used nomenclatures and methods in archetype concepts satisfying the Archetype Object Model of EN ISO 13606 [2][3]. In this register, the ICD 10 codes are mapped to the Orphanet nomenclature of rare diseases and respective ORPHA codes. For example, Figure 1 displays patient's names, sex, age and several anthropometric data - Height, Weight, and Body Mass Index. Records concerning the blood pressure include the Systolic and the Diastolic blood pressure, the method and the place of measurement and the technical details as type of device, position, level of the patient effort, inclination of the patient, and other conditions.

IV. CONCLUSION

This approach of standardization at all stages of data transmission gives the possibility to transfer the data between our national register and the international repositories, using the European standard for an International Patient Summary (IPS) (prEN 17269) [4].

ACKNOWLEDGMENT

This research is supported by the National Scientific Program "e-Health in Bulgaria".

REFERENCES

[1] D. Tcharaktchiev et al., MEDICA – 9 Years of Development and Use of a Clinical Information System in the University Hospital of Endocrinology and Gerontology – Sofia. in: Medical Informatics Europe '96, J. Brender et al. (Eds.), IOS Press, 1996, pp. 458-462.

[2] D, Tcharaktchiev, V. Gergiev, and I. Ivanov, "Archetypes and Standards for Medical Information Interchange". ACSIJ Advances in Computer Science: an International Journal, Vol. 4, Issue 5, No.17 , 2015, pp. 133-137.

Copyright (c) IARIA, 2020. ISBN: 978-1-61208-817-4

[3] CEN/TC 215, "ISO 13606-2:2019. Health informatics -- Electronic health record communication -- Part 2: Archetype interchange specification," ISO, 2019. [Online]. Available: https://www.iso.org/standard/62305.html.[last accessed: 22.10.2020].

[4] CEN/TC 251, "EN 17269 Health informatics - The Patient Summary for Unscheduled Cross-border Care". European Committee for Standardization, Brussels, 2018.

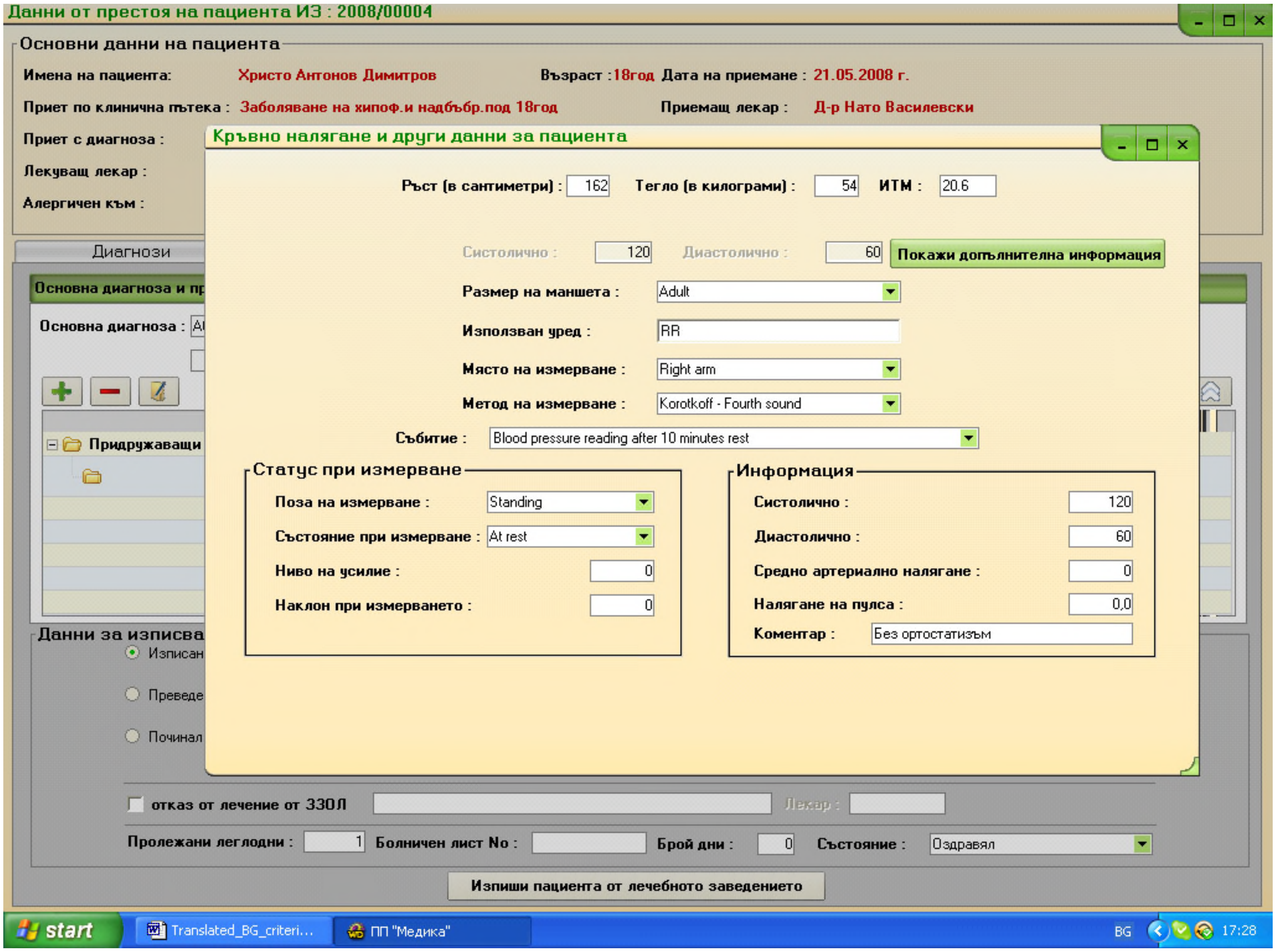

Figure 1. Blood Pressure Archetype (HIS of the University Specialised Hospital for Active Treatment of Endocrinology – Sofia, Bulgaria).

Management of Clinical Concepts in Bulgarian Healthcare Using openEHR Specifications

Simeon Abanos
Faculty of Mathematics and Informatics
Sofia University St. Kliment Ohridsky
Sofia, Bulgaria
email: simeonabanos@gmail.com

Evgeniy Krastev
Faculty of Mathematics and Informatics
Sofia University St. Kliment Ohridsky
Sofia, Bulgaria
email: eck@fmi.uni-sofia.bg

Dimitar Tcharaktchiev
Department of Medical Informatics
Medical University
Sofia, Bulgaria
email: dimitardt@gmail.com

***Abstract*—Clinical concepts in national healthcare are usually represented employing heterogeneous information models. The existence of incompatible information models significantly complicate interoperability and management of clinical documents across a country as well as cross-border exchange of such documents. The objective of this paper is to outline a methodology for management of typical clinical concepts in the scope of Bulgarian healthcare by means of openEHR archetypes. A case study of management involving widely used clinical concepts is here considered. The clinical concepts are designed with openEHR concepts and therefore they are ready for semantic interoperability. The obtained results prove the hypothesis that existing health documents like clinical path reports can be transformed into archetype object model by preserving the semantic context. New results from computer experiments are presented and discussed in the context of providing patient- centric eHealth services.**

Keywords-openEHR information model; openEHR archetype; openEHR server; clinical path report; semantic interoperability.

I. Introduction

Heterogeneous information models are widely used in national healthcare. Health information systems employing such models are incompatible. This complicates significantly management and declines the quality of eHealth patient services.

The objective of this paper is to outline a methodology for management of typical clinical concepts in the scope of Bulgarian healthcare by means of openEHR archetypes [1]. The methodology covers essential stages in the development a fully functional web application allowing semantic interoperability in management of clinical path reports by the National Health Insurance Fund (NHIF).

The following Section II shortly presents the proposed methodology for building an openEHR archetype model comprising typical clinical concepts in national healthcare. New results in the implementation of the archetype model are provided and discussed in Section III. Section IV makes an evaluation of the contributions in this research work in relation to results in the existing literature.

II. Methodology

This paper employs openEHR [2] information model concepts to build an archetype object model of a clinical path report according to the XML schema model provided by the NHIF. It has been successfully implemented in a related research project for cross- border exchange of International patient Summary standard [3].

The applicability of the algorithm is validated by processing a real-life collection of clinical reports on a openEHR server.

III. Results

An openEHR template of a clinical path report is designed. This report in national healthcare contains details of the medical services and procedure provided in the process of patient treatment. An algorithm is developed for transforming existing heterogeneous information models of clinical data in such reports into instances of a template composed of openEHR archetypes. The transformation procedure persists data structures and semantic context.

A multitier web application is developed for the purpose of testing the execution of typical management tasks with clinical concepts (Figure 1). It persists and manages archetype instances(contributions in openEHR server terms) of clinical path documents on openEHR servers[4]. These severs provide a common software platform and can be deployed locally in a healthcare organization or in a cloud infrastructure.

Clients are enabled to manage clinical path reports by invoking web service operations of the openEHR servers. In our computer experiments we demonstrate user friendly web interface that enhances the business process for openEHR management of such reports (Figure 2)[5]. This approach to modeling and management clinical data is being demonstrated for the first time in Bulgarian healthcare.

IV. Conclusion

The obtained results prove the hypothesis that existing health documents like clinical path reports can be transformed into archetype object model. The proposed information model enables a standards-based approach to management of clinical documents and facilitates semantic interoperability between information systems in healthcare.

Results from computer experiments demonstrate clinical path report data interactions among openEHR platforms.

This novelty approach in national healthcare allows finding cost efficient solutions. Its implementation extends our previous research work carried out in developing solutions for patient-centric eHealth services.

Acknowledgment

This research is supported by the National Scientific Program “e-Health in Bulgaria”.

References

[1] openEHR, Archetype Model (AM) Component - latest. 2020. [Online] Available at: https://specifications.openehr.org/releases/AM/latest [retrieved: October, 2020].

[2] openEHR, Reference Model (RM) Component - latest. 2020. [Online] Available at: https://specifications.openehr.org/releases/RM/latest [retrieved: October, 2020].

[3] E. Krastev, D. Tcharaktchiev, P. Kovatchev, and S. Abanos, "International Patient Summary Standard Based on Archetype Concepts" International Journal on Advances in Life Sciences, ISSN 1942-2660, vol. 12, No. 1 & 2, year 2020, pp. 34-46, http://www.iariajournals.org/life_sciences/ .

[4] P. P. Gutiérrez, "Towards the Implementation of an openEHR-based Open Source EHR Platform (a vision paper)", Studies in Health Informatics, vol. 216, pp.45-9, 2015.

[5] P. Kovatchev, "Convert and insert hospital NHIF records into openEHR EHRServer instance," 2019. [Online]. Available: https://www.youtube.com/watch?v=a6bWEog98U4&ab_channel=PetkoKovachev[retrieved: October, 2020].

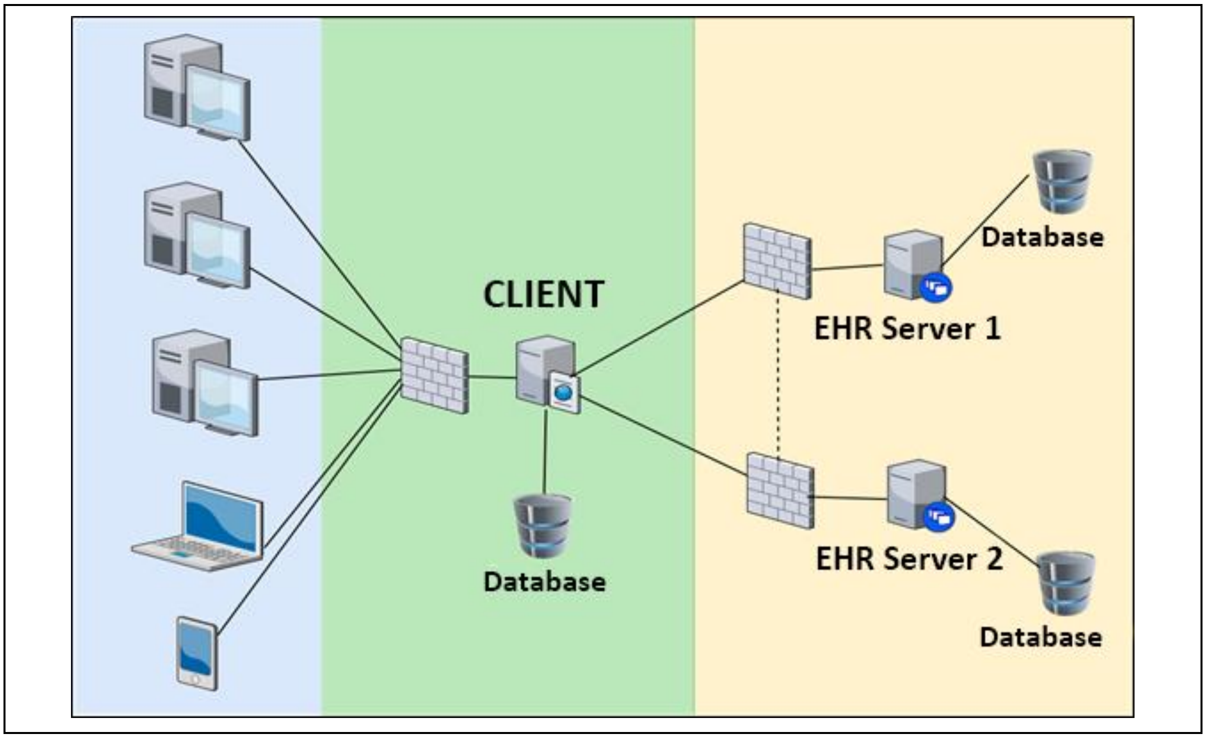

Figure 1. Architecture of the multitier web application for magement of clinical concepts.

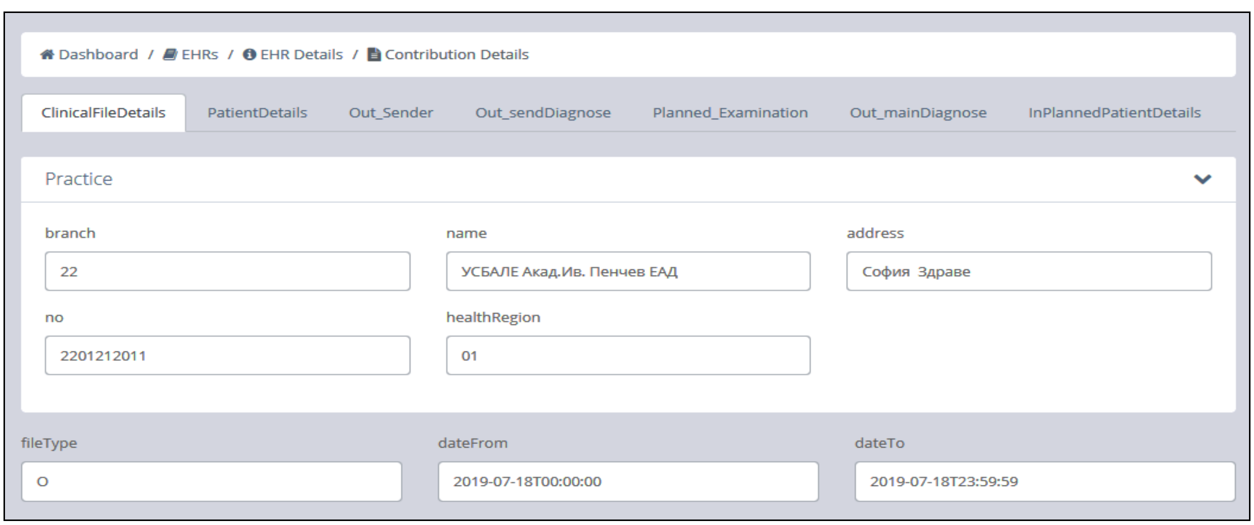

Figure 2. Web interface for management of clinical concepts..

Modelling and Management of ePrescriptions on openEHR platform in Bulgarian eHealth

Petko Kovachev
Faculty of Mathematics and Informatics
Sofia University St. Kliment Ohridsky
Sofia, Bulgaria
email: az@petko.info

Evgeniy Krastev
Faculty of Mathematics and Informatics
Sofia University St. Kliment Ohridsky
Sofia, Bulgaria
email: eck@fmi.uni-sofia.bg

***Abstract*—Medication prescriptions are widely used in patient-centric eHealth services. In national healthcare they are usually represented employing proprietary information models. This significantly complicates management of prescriptions in electronic format (ePrescription). The lack of interoperability among such ePrescription models increases both the risks of drug prescription errors and the costs for ePrescription processing. The objective of this paper is to outline a methodology for management of ePrescriptions in the scope of Bulgarian healthcare by means of openEHR archetypes. The case study in this research work considers a typical use case in this application domain. A widely used information model of a medication prescription is represented in terms of reusable openEHR archetypes. The use of openEHR entails semantic interoperability and portability of ePrescription instances across health information systems. Results from computer experiments reveal ePrescription benefits like reduced medication errors, increased patient safety and improved management of electronic health records.**

Keywords-openEHR information model; archetype object model; ePrescription; openEHR platform.

I. INTRODUCTION

The need for improved quality of health services is one of the main reasons for adopting prescriptions in electronic format (ePrescription) as a preferred way for purchasing medicinal products [1]. The objective of this paper is proposing a methodology for modelling and management of ePrescriptions using openEHR specifications [2]. The following Section II shortly presents the proposed methodology for building an openEHR archetype model comprising typical clinical concepts in national healthcare. The status of deployment and the challenges in the implementation of the proposed approach are presented in Section III. Section IV makes an evaluation of the contributions in this research work in relation to results in the existing literature.

II. METHODOLOGY

The proposed ePrescription model is an openEHR template composed of archetypes corresponding to clinical concepts in existing National Health Insurance Fund (NHIF) prescriptions.

This approach reduces significantly the difficulty and cost in application development. Moreover, it ensures semantic interoperability, reusability of components and portability of patient-centric services across systems.

III. RESULTS

An open source openEHR platform is configured for evaluation the practical implementation of basic functional requirements of the model [3]. The obtained results extend previous research work in the area of cross-border exchange of International Patient Summary [4]. An INSTRUCTION archetype model of ePrescription is created employing openEHR specifications (Figure 1). The model incorporates all the health data in a prescription as it is specified by the NHIF. A multitier web application is developed for database management of ePrescriptions on openEHR platform [2]. The client part of the application communicates with the openEHR platform by means of web services (Figure 2). It is used to create and retrieve ePrescriptions with real-life data employing user friendly web interface. The design of the user friendly web interface enhances the business process for openEHR management of such reports. The information model is implemented in a fully functional web application

IV. CONCLUSION

The here proposed methodology allows to transform existing XML schema definitions used by the Bulgarian NHIF into archetype-based models with inherent semantic interoperability of clinical documents exchanged by other EU countries. The implementation of this methodology is novel in the existing literature, where known ePrescription software applications allow basic functional interoperability of clinical documents. Plans for future results include the development of an application for cross-border exchange of ePrescription that add semantic interoperability to current “epSOS- friendly” ePrescription documents.

ACKNOWLEDGMENT

This research is supported by the National Scientific Program “e-Health in Bulgaria”.

REFERENCES

[1] C. Coimbra et al., "Improving the Codification of Hospital Discharges with an ICD-9-CM Single-page Application and its Transition to ICD-10-CM/PCS," International Journal on Advances in Life Sciences, vol. 10, no. 1&2, pp. 23 - 30, 2018

[2] openEHR, Reference Model (RM) Component - latest. 2020. [Online] https://specifications.openehr.org/releases/RM/latest [retrieved: October, 2020].

[3] P. Kovatchev, "Web based software for converting NHIF claim XML health records to openEHR specification," 2020. [Online] https://www.youtube.com/watch?v=YQoRCESMt [retrieved: October, 2020].

[4] E. Krastev, D. Tcharaktchiev, P. Kovatchev, and S. Abanos, "International Patient Summary Standard Based on Archetype Concepts" International Journal on Advances in Life Sciences, ISSN 1942-2660vol. 12, no. 1 & 2, year 2020, pp. 34-46 [Online] http://www.iariajournals.org/life_sciences/ [retrieved: October, 2020].

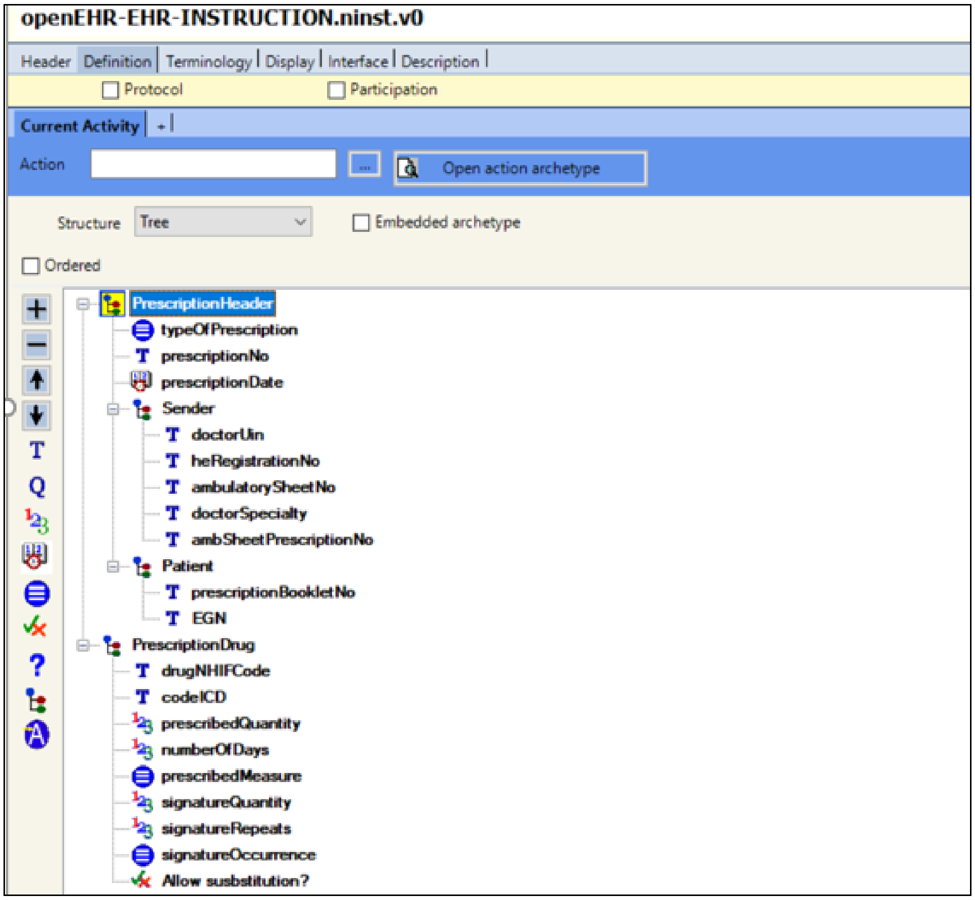

Figure 1. openEHR INSTRUCTION archetype with ePrescription clinical data.

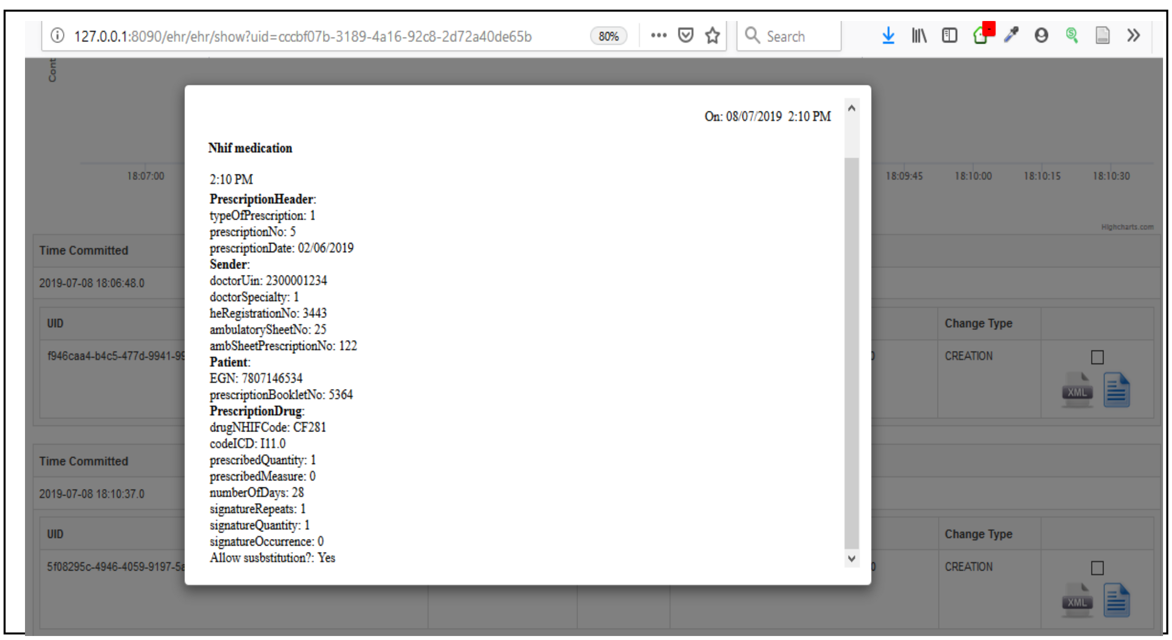

Figure 2. Extraction and visualization of ePrescripton data in openEHR server's web console.

Medical Requirements for Selecting Local Picture Archiving and Communications Systems

Influences from Information Technologies and Business Models

Ivan Evg. Ivanov
Technical University
Sofia, Bulgaria
email: iei@tu-sofia.bg

Dimitar Tcharaktchiev
Medical University
Department of Medical Informatics, USHATE
Sofia, Bulgaria
email: dimitardt@gmail.com

Vesselin Gueorguiev
Technical University
Sofia, Bulgaria
email: veg@tu-sofia.bg

Desislava Georgieva
New Bulgarian University
Sofia, Bulgaria
email: dvelcheva@nbu.bg

***Abstract*—For a period of 10-15 years (2005 - 2018), the number of PACS (Picture Archiving and Communication System) offered on the market has increased more than 10 times. Despite this growth, some of the main disadvantages of this class of systems persist: the systems are complex and expensive to acquire, replace, maintain, or repair. This paper aims to show an approach that allows us to choose the right PACS for the needs of a particular group of hospitals and health care institutions. This choice takes into account the limitations of IT and business requirements.**

Keywords- regional (local) PACS; PACS selection; regulatory requirements; business requirements; medical requirements.

I. INTRODUCTION

The last two decades have seen a steady increase in the clinical application Picture Archiving and Communication Systems (PACSs) [1]. The reason for this increase is the desire to both reduce costs and improve patient care. At present, this is a sustainable trend, which suggests its significant impact on the speed of changes in medical infrastructure. The problem is to select PACS for long-term use that is adaptive for extensions, user-friendly for medical personnel and easy-to-connect to other medical institutions, upper-level PACSs and medical information systems.

II. PACS MAIN COMPONENTS AND STRUCTURE

The history of the development of PACS goes through several stages. Perhaps the most characteristic of all these stages was that the systems were oriented towards the practically identical infrastructure architectures [1]-[4]. This trend of the regularity of the infrastructure is preserved, notwithstanding the increase in the use of newer programming technologies and newer types of medical devices. One of the main features for comparing modern PACSs is the ability to adapt to the specific infrastructure, i.e., to the possibility of removing or adding new PACS nodes or communication buses.

The general infrastructure of modern PACS contains the following elements (see Figure 1):

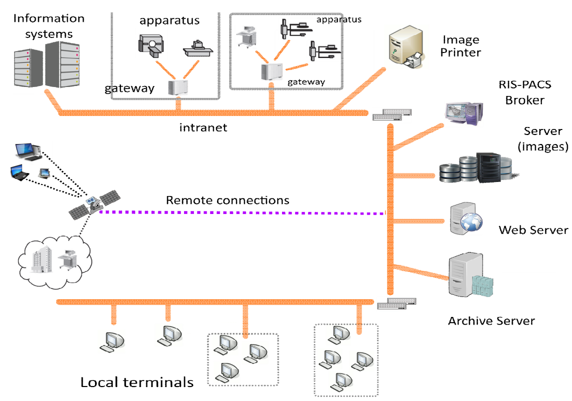

Figure 1. The general infrastructure of modern PACS (adapted from [4]).

- Apparatus that generates medical images in an appropriate format (most often, PACS uses one or more versions of Digital Imaging and Communications in Medicine (DICOM).
- An image storage server. This is the PACS' key component.
- The archiving system is the second key component of a PACS.
- The interoperability brokers responsible for solving the task of integrating the PACS with the other information systems in the health organization.
- The terminal workstation (i.e., the computer system for displaying medical images).
- The printer for medical images.
- Remote connection to PACS. Now, it is mandatory function for any more or less "professional" product in this group.

III. PACS: KEY BUSINESS REQUIREMENTS

The need to use medical images throughout all hospital departments, as well as the possibility of their remote use by authorized external organization or physician, has defined the following list of mandatory business requirements for modern PACS:

Copyright (c) IARIA, 2020. ISBN: 978-1-61208-817-4

- Any strategy for the management and dissemination of medical images within the hospital must provide the opportunity for a wide variety of medical users to work effectively. These are not just physicians in functional radiology departments. Therefore, the system must not change physician's stereotypes of working with images but at the same time must allow managing the image quality according to needs of the diagnostic process.
- No one likes to wait to get data from some source. Clinicians are not an exception to this rule. Very often, they are even more demanding on the access time and restrictions to obtain the whole required information about the patient. The most common manifestation is the desire for almost immediate access and it does not depend on whether physicians work in traumatology, in the emergency rooms, in the intensive care department, or the outpatient clinic. This business requirement is one of the most complex problems in everyday life.
- Improving the ability to diagnose and treat many diseases has increased the use of medical images in practice. From a technical point of view, this change has led to increased size of generated images, increased number of requests to use medical images, and the size of archived history of patient morbidity needed to monitor the condition of patients. All this changed the understanding of the minimum required levels of scalability, efficiency, performance, and adaptability of the system.
- In the diagnostic process, to avoid mistakes, physicians need the best possible image quality. A very serious technical problem is the fact that the judgment of quality varies according to the particular usage of the image. The only workable solution to this problem is the presence of a system for automatic or semi-manual adaptation of the display workstation to environmental characteristics. This complicates the PACS functions responsible for handling workstations' settings.
- Every doctor has his understanding of the workstation's ergonomics. The main differences are not only in the arrangement of buttons and menus on the screen, but the scenarios of using images and meta-information in the diagnostic process
- As part of daily activities, physicians integrate information from medical images with other clinical data to make decisions about patient's diagnosis and treatment. Much of this additional information is not directly available and needs to be obtained from the information systems of other hospital departments. The main problem, in this case, is the compatibility of data storage and data transfer standards.
- Physicians not only use data from patient's medical examination, but also they generate new data about the patient's medical condition. Therefore, PACS needs to provide an opportunity not only to review existing data but also to add new data to the patient's history file. The complexity of the problem is not in the data insertion itself but in the ability to validate the added data and to verify the process of data insertion.
- Medical systems store and handle patients' data. The solution to the problem is to use very high-level standards of data security. The extension of a PACS with remote access functionality (e.g., Internet-based remote connections) increases the complexity of the problem.

Regardless of the existing rules and regulatory provisions, when a new PAC has to be installed in a hospital, it is mandatory to answer the question of how these requirements are implemented. As not all requirements are always met, the system is considered allowed if there is a mechanism to add the necessary functionality.

IV. CONCLUSION

A PACS is purchased and installed with an intention of long time use without replacement. At the same time, like most computer systems in this category, PACS has a complex structure, internal interactions, and implementation. Frequently occurring problems can create significant inconveniences at work, including blocking routine clinical activities due to the unavailability of PACS resources or functions. This constitutes a great complexity of static approaches to new PACS characteristics validation.

Therefore, we explored another way to pre-assess the suitability of a PACS for an individually selected hospital or group of hospitals.

The article represents our approach, which optimizes costs and time for selecting a new PACS for installation and long-term use. The main phases of the approach are outlined, as well as the required activities for each of the phases. The presented approach is put in practice in a government-funded project for a group of hospitals.

ACKNOWLEDGMENT

This research is supported by the National Scientific Program “e-Health in Bulgaria” D01-200/2018.

REFERENCES

[1] H. K. Huang, PACS: Picture Archiving and Communication Systems in Biomedical Imaging, VCH, 1996, ISBN: 978-1560816850

[2] H. K. Huang, PACS-Based Multimedia Imaging Informatics: Basic Principles and Applications, 3rd Edition, Wiley-Blackwell, 2019, ISBN: 978-1118795736

[3] A. Peck, Clark's Essential PACS, RIS and Imaging Informatics, CRC Press, 2019, ISBN: 978-1498763233

[4] E. Samei et al., "AAPM/RSNA tutorial on equipment selection: PACS equipment overview: general guidelines for purchasing and acceptance testing of PACS equipment" in Radiographics, Jan-Feb 2004; 24(1), pp. 313-334, doi: 10.1148/rg.241035137.

Copyright (c) IARIA, 2020.

Cancer Caregivers' Needs for Their Well-Being in the Information Era

Koo Ah-Choo, Peter Charles Woods, Tenku Putri Norishah Binti Tenku Shariman
Faculty of Creative Multimedia
Multimedia University
Cyberjaya, Malaysia
e-mail: ackoo@mmu.edu.my, p.woods@mmu.edu.my, tengku.norishah@mmu.edu.my

Teoh Sian-Hoon
Faculty of Education
Universiti Teknologi MARA
Puncak Alam, Selangor
e-mail: teohsian@salam.uitm.edu.my

Ang Kok-Yew
Department of Game Studies, Faculty of Creative Industries,
Sungai Long, Kajang,
Universiti Tunku Abdul Rahman e: angky@utar.edu.my

Chin Weng-Ping
Education Volunteer
Puchong Hartamas, Puchong, Selangor
e-mail: lydiadigitalclassroom@gmail.com

***Abstract*— Caregivers face numerous challenges when dealing with chronic diseases such as cancer. Ageing populations and environmental factors account for increasing numbers of people suffering a variety of cancers, and inevitably the increase in caregivers. Caregivers may be the influencers for cancer treatment decisions by patients, hence their needs as caregivers warrant study. This paper focuses on identifying factors of cancer caregivers' needs in the context of an information and communication technology oriented society. A total of 84 caregivers were recruited for the study from two cancer clinics in a hospital, located in Kuala Lumpur. Quantitative research using a survey instrument was conducted. Factor Analysis was used to identify the need factors, that are communication, personal well-being, basic healthcare, information access, coping with change, and learning needs. This study identified the need for improving communication between cancer sufferers and caregivers, from the caregivers' perspective is seen to be crucial. Facilitation on communication needs should be incorporated into any patient-centered healthcare system for establishing a comprehensive healthcare model. Future research is proposed to look more specifically into information, communication and learning needs.**

Keywords- Cancer; caregivers' need; communication; information; factor analysis.

I. INTRODUCTION

Being in the socially connected and information era, caregivers face challenges in selecting and using information for decision making in their caregiving tasks and decisions. Reference [29] identified the importance of health (cancer) literacy for caregivers. They mentioned that "accessing information from the internet may entail additional demands and capacities compared to traditional health literacy due to the factor of competing sources, identifying accurate and trustworthy resources, technological and internet literacy, as well as access to technology and the internet" (p.12). They propose future research could examine various strategies in providing information to caregivers through eHealth modalities. Reference [18] suggested that innovative health education such as mobile learning applications will further expand the context of a smart learning ecosystem for cancer education.

Caregiver is defined as "someone who performs hands-on care and/or provides emotional support to patients, such as a partner, relative or friend" ([6], p. 388). "Formal caregivers typically undergo training and certification, and may inherently have greater health literacy capacities compared to informal caregivers" ([29], p. 12). The term "caregivers" used in this study is to encompass both partners and family caregivers, who are inherently informal caregivers.

Caregivers to elderly recipients have a significant influence on their treatment [17]. Support is needed for caregivers, such as physical, psychological, social, and spiritual [1] [20] and encompass many decades of care for patients with chronic neurological conditions such as multiple sclerosis, Alzheimer's, Huntington's disease, stroke and Parkinson's disease [22]. However, the capacity of informal caregivers to source and utilize information in order for them to perform their tasks is not well understood [30], especially in the current socially connected world. Caregiving tasks are complex. Increasingly, the family and society are aging and the chance of elderly affected by the disease is higher. The roles of family caregivers are multifaceted and understanding their needs are crucial.

The purpose of conducting this study is mainly to identify the cancer caregivers' needs within the context of information seeking and informal learning model. The need factors were adapted based on [6], with additional factors on informal learning in the information era as highlighted in the literature review. The study reported a detailed process of Principal Component Analysis (PCA). The identified need factors may contribute to the model of cancer caregivers' needs in the information era.

In Section II and its sub-sections, the relevant literature is reviewed on the aspects of dealing with cancer, information access on caregiving and cancer caregivers' needs. The rest of the sections report the methods, analysis and findings, discussions, conclusion and future works of the research.

Copyright (c) IARIA, 2020.

II. LITERATURE REVIEW

A. *Dealing with cancer*

Annual cancer cases globally are expected to rise from 14 million in 2012 to 22 million within the next 2 decades [26]. Cancer rates increase according to age [16]. A total of 18,219 new cancer cases were diagnosed in 2007 and registered with the Malaysia National Cancer Registry, with 8,123 (44.6%) males and 10,096 (55.4%) females [16].

Giving care to cancer patients is a challenging task for family caregivers [2] [20]. Informal caregivers are untrained [22]. This affects the quality of care in a patient-centred care model, especially when the care is carried out at home without the guidance or support from healthcare professionals. There is a need to increase the awareness of cancer care for caregivers.

According to the President and Medical Director of National Cancer Society Malaysia (NCSM), Dr Saunthari Somasundaram, there is a need to provide awareness about the misconceptions and myths around cancer in Malaysia's culture, and requires a 'whole-of-society' approach to tackle this issue [3] [10]. Prompt actions need to be taken on the misconception to decrease the risk of cancer, and detect cancer earlier for faster access to treatments. Reference [8] urged that the cancer care system in Malaysia requires an urgent reform as their study found that "…Out of the 2,312 deaths due to breast cancer, 2,048 (88%) were avoidable. Of these avoidable deaths, 1,167 (57%) were attributable to late stage presentation while 881 (43%) were due to lack of access to optimal treatment" (p.32). This statistic shows an under-performing outcome of the cancer care system in Malaysia. Reference [8] further identified a Malaysian cancer divide between the rich and poor, mirroring the global cancer divide between rich and poor countries. They said, "…the huge number of avoidable deaths highlights the high cancer mortality rate among the deprived and the vast disparity in access to cancer care between the rich and poor within Malaysia." (p.32)

Dealing with the cancer demands both the management of the disease and its consequences emphasise that "cancer is a societal issue and is not confined to health" [28], which undoubtedly needs more attention in research. The social groups for cancer support could play a more proactive and important role to promote early treatment or disseminate accurate information about cancer. There is also a need for palliative care to reduce caregivers' burden [1] [4]. Compared to caregivers of other types of illnesses, cancer caregivers spend more hours providing care and the intensity of their caregiving is increasing due to the nature of active treatments. They are definitely in need of assistance and information to make the best decision regarding cancer treatments [11]. The length of time spent on caregiving determines the stress levels among the caregivers. Access to caregiving support in the society or community is often limited in Asian countries. According to [28], support was mainly acquired through non-governmental organizations (NGOs). Access to integrated care that extends from hospitals to the community is lacking, especially for healthcare needs, and trustable informational needs [28]. Access to information is mostly informal, influenced by many viewpoints, ranging from information provided by medical professionals, to information retrieved from the internet, friends, family, and social media sites that are often conditioned by the belief system grounded within the various cultural groups. These multiple views clearly pose a challenge for patients to make wise treatment decisions. It is the responsibility of the government, civil society and communities to support the process and education of patients and caregivers in dealing with cancer. Healthcare expenses have risen in recent years, hence cancer care is a financial burden for most households and the country [13]. Treatment costs in the first year after diagnosis can exceed 30% of household income in ASEAN countries [13] [28].

B. *Information access on caregiving: A mixed blessing*

We acknowledge the idea of a Seamless Learning Model in the context of caregivers' informal learning and information seeking pursuits, implemented seamlessly via the internet and social media, which impacts their public and private learning spaces. Many activities of learning and information seeking, especially for caregivers are inherently informal, self-directed, independent, and critical as they are frequently influenced by online technologies and social media. There are many internet-based information platforms for supporting and developing skills in caregiving and social care, for example [33] [34] [35]. It is also recognized as an authentic and just-in-time learning (or training), especially for caregivers or informal caregivers who need some help and guidance during challenging times.

Many countries have developed their national health websites and awareness programs to spread knowledge about healthcare and promote a healthy lifestyle by using the internet and social media platforms (e.g., Myhealth Portal and Infosihat for Malaysia; Singapore Health Promotion website). However, according to [2] (p. 627), the importance of these online media and technologies has not been clearly revealed in previous studies, especially on the information-seeking behavior of family caregivers. More research should be conducted to add literature on the information seeking behaviours among caregivers.

Through seamless technologies, [21] propose the use of tools to design diverse learning experiences such as creating extended learning communities, linking people in real and virtual worlds (connectedness), providing expert knowledge based on demand, and to supporting learners in many ways. Reference [14] proposed a Seamless Learning Framework (Figure 1) to explain the learning environments or dimensions experienced by most of the learners. The learning space is no longer defined by a "physical / formal class" but by "learning unconstrained by scheduled class hours or specific locations" (p.156), thus promoting seamlessness, with informal learning and information access at learners' fingertips. The seamless environment is labeled as "community" which comprises different categories of people such as teachers, experts and learners. The community has access to any relevant sources of knowledge through cognitive tools, within the dimension of time (anytime), space (anywhere), and artefacts (any learning artifacts / contents).

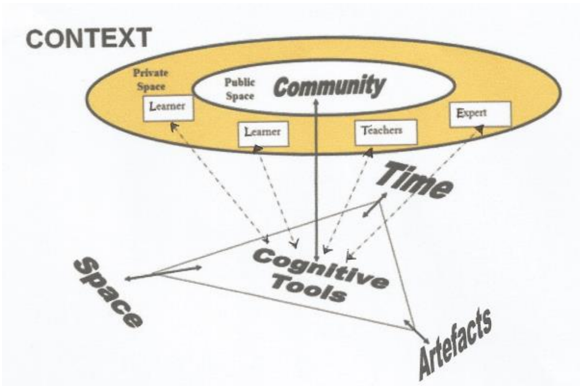

Figure 1. A Seamless Learning Framework [14].

This concept can be applied to fulfil the informal learning needs of cancer caregivers since caregivers interact with all kinds of information and materials seamlessly. They interact with internet support groups or cancer survivors through social media; and also conduct discovery learning about the disease from the cancer journeys shared by others. There are also professional cancer learning and sharing materials which can be accessed with ease.

The internet is a 'mixed blessing' for the healthcare sector. Reference [25] suggested that public health information campaigns could be conducted using social media. This new means of communication, especially for prevention purposes, will complement other methods of communication. However, the challenge will be to adhere to the legal framework that preserves the quality of the healthcare information provided on the internet.

Consumption of online healthcare information and services is increasing [15] and has empowered caregivers (and patients) to enhance their health literacy for improved decision making [9].

Reference [18] conducted a quasi-experimental effectiveness study of a new mobile learning tool for cancer education and the results showed that the application has significantly increased cancer prevention knowledge among the users. However, the main drawback of online healthcare information is its credibility. Diagnostic information online is sometimes used to confront doctors [12] [15]. Reference [15] suggested that users, such as caregivers must critically analyze and select information posted on YouTube in order to make effective healthcare decisions.

A. *Cancer caregivers' need*

Reference [19] - Seattle Cancer Care Alliance (2013) listed the challenges of caregiving: caring for themselves (self-care), supporting and caring patients emotionally and physically, maintaining the home environment for patients, gathering information, helping with decision-making on cancer care or treatment, arranging patients' hospital visits and patient's financial support. These tasks are complex especially for the new caregivers. Roles of family caregivers are multifaceted, and adequate support is required for them (Glajchen 2004 in [2]). According to Wingate and Lackey cited in [2], family caregivers need knowledge, information, or understanding that can be gained through education, experience, study, or through explanations by qualified cancer specialists. Reference [2] reported a qualitative study that discloses the information needs of cancer family caregivers are varied along the cancer journey, and they used diverse information sources, including healthcare professionals, hospital booklets, interpersonal networks, besides the internet, mass media, and books - to satisfy their needs. Her study found that demographic variables of caregivers (such as gender, age, level of education, socioeconomic status and culture) affected their information-seeking behaviors.

Reference [6] measured the psychometric properties to capture the multidimensional supportive care needs of cancer caregivers. The instrument used was given to 547 cancer caregivers. Psychometric analyses found four dimensions of need: healthcare service, psychological and emotional, work and social, and information. Individuals with anxiety and depression were more likely to report at least one unmet moderate or high need in comparison to non-anxious participants. Younger participants experienced at least one unmet moderate or high need in the area of "psychological and emotional", and "work and social", as compared to the older participants. Reference [6] suggested the findings can be used to prioritize healthcare resources and tailor supportive cancer care service accordingly.

In summary, the study intends to explore the properties or dimensions of needs faced by caregivers, especially the four dimensions namely healthcare service, psychological & emotional, work & social, and the last dimension is the information need. Additional dimension of learning need is also an interest of the current study.

III. METHODS

A survey was conducted in a hospital with two cancer specialist clinics run by NCSM, an NGO for cancer awareness and cancer care. Permission was granted from NCSM in order to conduct this study at the waiting lounge of the two cancer clinics. The population of the study involved all cancer clinics run by NCSM across the country. However, only two cancer clinics were purposively selected to participate in this study due to their strategic location in the center of Kuala Lumpur city. A total of 84 participants were involved in this study. Majority of them were Malaysian Chinese. This is because the hospital is traditionally or historically relevant and popular among Malaysian Chinese community. In addition, the Chinese in Malaysia have the higher lifetime risk for cancer compared to Indians and Malays. Chinese is the second largest ethnic group in Malaysia who form approximately 24 percent of the Malaysian population [16].

Previous reports of the National Cancer Registry (NCR) [16] have revealed that cancer seems to be predominant among Chinese as compared to Malay and Indian. "The Age-Standard-Rate (ASR - a measure for cancer incident) for Chinese male was 111.9 per 100,000 and for Chinese females was 115.0 per 100,000 while for Indian male was 68.2 per 100,000 populations and for females was 99.9 per 100,000 populations." [16] (p. 23). According to the same report, "the

Copyright (c) IARIA, 2020. ISBN: 978-1-61208-817-4

ASR among Malay male and female were 66.9 per 100,000 and 79.0 per 100,000 populations respectively." (p. 23)

A. Instrument

The instrument of the survey was adapted from [6] to study needs in the context of informal learning environments. Items related to this context were added to the instrument. The instrument was then reviewed by three experts in the area of wellness and preventive medicine (Expert 1), management of a cancer wellness center (Expert 2), and healthcare informatics (Expert 3) respectively. Items were reviewed; some were dropped and merged because of overlapping meaning; some were rephrased for the suitability of local caregivers in Malay and English.

The items were analysed using factor analysis. Factors on the needs of cancer caregivers were then identified in the analysis. There are two key sections of the questionnaire:

Section A – Profiles, caregiving experiences and online activities behavior, Section B – Needs measured by Likert-type items, adapted mostly from [6], that is ranging from "1 denotes no need (either satisfy or non-applicable), 2 denotes low need, 3 denotes moderate need, and 4 denotes high need".

B. Procedures

The caregivers in the waiting lounge were invited to participate in the survey. The researcher assistant took turns to be in the waiting area for two or three days in a week for two months. The data collection process stopped when there were very few new caregivers in the waiting area.

Caregivers who agreed to participate in this study were asked to sign a participation consent form. They were briefed that at any time, they could freely withdraw from the study. Some caregivers were assisted by the researchers to clarify the meaning of items. Overall, the total number of usable responses was 84 out of 91; seven responses were not included due to incompleteness.

IV. ANALYSIS AND FINDINGS

There are two main stages of factor analysis conducted for extracting and determining factors in this study. Firstly, factor extraction was performed focusing on making initial decisions about the number of factors that primarily satisfies a set of measured variables based on the Principal Component Analysis (PCA). Secondly, factor rotation was conducted to ensure the underlying factors were more interpretable.

A. Description of the demographic information

More than half were females, the majority of them were Chinese (83%) with their religion Buddhism or Taoism, middle income or lower (77%), holding diploma and above (59%), aged more than 30 years old (73%). More than half of them (62%) have been caregiving for 6 months and above, mostly the care recipients were affected with female related cancer (54%). The caregivers were quite active online to look for health information and connection with others via online. About 60% of them were in this category.

B. The steps of factor analysis

Factor analysis was employed to establish factors of needs, through extracting factors and factor rotation procedures as recommended [24]. Principal Component Analysis (PCA) is chosen because it is a common method for exploring loadings of survey items into the components and it is frequently used in research of education and psychology [23] [24]. The PCA allows transformation of a set of inter-correlated variables into specific components or factors to extract distinct factors from the variables in the analysis which involves reduction of dimensionality among the variables. In the process of extracting factors, suitability of data was tested, and hence actual component extraction was retrieved for getting a smaller number of items to best represent relationships among those items.

1) Stage one: Extracting factor

In PCA, the suitability and adequacy of data in terms of variability of data were tested based on Kaiser-Mayer-Olkin (KMO) measure of sampling adequacy. In this study the value of KMO is 0.842 which is greater than 0.7 which indicates a very good condition to proceed with the factor analysis. In normal practice, the value of KMO should be larger than 0.5 for achieving a condition of satisfactory.

On the other hand, the Bartlett Test of Sphericity takes consideration of testing of correlations among the variables. This value is referred to ensure that there are sufficient correlations among the variables. The sufficiency of correlations is indicated in the associated probability in the chi-square. If the p value of the associated chi-square statistic is less than 0.5, it shows the items are sufficiently correlated for further analysis in PCA. In this study, the results of Bartlett Test of Sphericity show that value of chi-square is 2470 with df=703 and p-value < 0.05, indicating that the variables were sufficiently correlated to form the specific components and factors.

TABLE 1. KMO AND BARTLETT'S TEST.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.842
Bartlett's Test of Sphericity	Approx. Chi-Square	2470.094
	Df	703
	Sig.	.000

The number of extracted factors or components is determined from variances explained among the variables. The factor extraction also depends on eigenvalues. An eigenvalue is the variance or the variables accounted for by a factor. The six components which have eigenvalues more than one (namely between 1.187 $\approx$1.2 to 16.714$\approx$16.72) were selected into the factors. These values provide a guide to ensure that the variances of components are all contributed from more than a single observed variable (item).

In this study (Table 2, refer to the last page of this paper), the examination of the total variance explained by the components extracted were displayed in Table 2: An analysis of the total variance explained by components (refer to the last page of this article). Component 1 shows a very high

Copyright (c) IARIA, 2020. ISBN: 978-1-61208-817-4

eigenvalue and a high percentage (14.57%) of variance explained. Component 2 until component 6 has the total variance explained value of greater than 3% (both for Initial Eigenvalue column and other columns) were chosen. Other components have lower value of initial eigenvalue, which are lower than 1.187 (or its initial percentage of variance explained, which are lower than 3%, i.e. the bold number) are not planned to be selected for the next step of analysis. Thus, the six factors are explained with the total variance 64.1%.

An examination of the scree plot of the Eigenvalue versus Component has shown a clear "knee point / elbow", that is at the point of component number of 6 (refer to Figure 2). Other components have lower eigenvalue, which will not be considered for the next step.

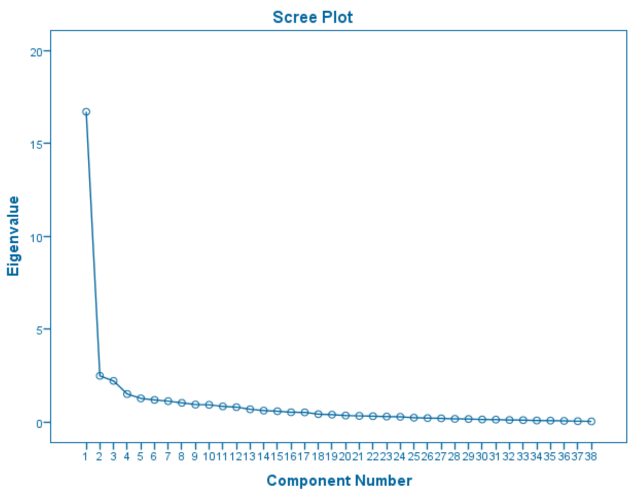

Figure 2. Scree plot showing the elbow at the point of the sixth component at the value of Eigenvalue >= 1.

2) Stage two: Factor rotation

The final step of PCA is to examine the factor loadings of each item in relation to the first six components. The factor loading presents the results of component rotations and interpretation of components.

The factor rotation method used in the analysis is Varimax rotation, a method used frequently in social science and psychological study. A check on oblique based rotation has also produced the similar set of items according to these components. In determining the factors from the factor loading, the loading of absolutes 0.4 is used as a cut-off value [7]. For practical significance, loadings of absolutes0.4 and above, but less than 0.5 are considered just enough to be significant. Loadings above 0.5 indicate highly significant. On the other hand, if the factor loading difference between two factors across a particular item is less than 0.2, the item should be dropped.

Table 3 is a matrix table showing the rotated components and their related items shows the loading value of each item which are greater than 0.4 in relation to the six identified components in the factor extraction, and other components (i.e. Components 7 and 8). The bold loading values are the selected items according to components; items which loading values are *italicized* due to the difference of loading values less than 0.2 were discarded (i.e. Item 18, 35, 26, 33 and 32). Components 7 and 8 were also not considered for interpretation.

TABLE 3. A MATRIX TABLE SHOWING THE ROTATED COMPONENTS AND THEIR RELATED ITEMS WITH FACTOR LOADINGS.

Item #	1	2	3	4	5	6	7	8	Difference of two factor loadings with nearest values	Decision on item#
14	0.819									
13	0.809									
19	0.658									
29	0.641									
20	0.622									
18	*0.575*				*0.433*				0.142 (<0.2)	Item 18 is dropped from analysis
15	0.469									
30		0.769								
27		0.766								
36		0.701								
28		**0.679**			**0.456**				**0.223 (>0.2)**	**Item 28 is maintained**
25		**0.638**			**0.424**				**0.214 (>0.2)**	**Item 25 is maintained**
10		0.611								
35		*0.593*	*0.467*						0.126 (<0.2)	Item 35 is dropped from analysis
26		*0.546*			*0.448*				0.098 (<0.2)	Item 26 is dropped from analysis
9			0.656							
6			0.605							
7			0.54							
33	*0.411*		*0.518*			*0.426*			0.015 (<0.2)	Item 33 is dropped from analysis
32			*0.515*		*0.486*		*0.432*		0.054 (<0.2)	Item 32 is dropped from analysis
8			0.476							
11			0.462							
5			0.45							
12			0.44							
3				0.783						
1				0.781						
2				0.706						
4			**0.425**	**0.64**					**0.215 (>0.2)**	**Item 4 is maintained**
24					0.787					
23					0.732					
17					0.506					
37						0.814				
38						0.809				

Copyright (c) IARIA, 2020. ISBN: 978-1-61208-817-4

34						0.594				
21							0.757			
31			*0.429*				*0.444*		0.015 (<0.2)	Item 31 is dropped from analysis
22								0.624		
16					*0.418*			*0.5*	0.082 (<0.2)	Item 16 is dropped from analysis

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 10 iterations.

Note: The minimum loading score accepted is 0.4, loadings below 0.4 are not shown; *items with loading difference <0.2 are discarded from further analysis / interpretation*. Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. A. Rotation converged in 10 iterations.

The bold items were retained and used for interpreting the need factor. In total, there were 29 items to explain six (6) components of need factors. Table 3 shows a matrix of the rotated components and their related items. A total of 9 items were discarded, and not included for further interpretation of the construct or need factors.

The internal reliability analyses with the Cronbach's Alpha values for the items emerged for six components / factors are stated in column 3 of Table 4: Factors with items loaded to the six-factor of needs (Refer to the last page of this paper). All factors have the alpha values greater than 0.7 ($\alpha >$ 0.7), indicating an acceptable internal reliability measure for the factors.

Collectively, none of the factors scored 'high need' and there is no factor indicating "no need". The need for communication is at the level of 'moderate'. Other need factors such as personal well-being, basic healthcare, access to information, coping with change are between 'low' to 'moderate need'. The need for learning through online information and connection is 'low'.

V. DISCUSSION

Six-factor of needs were identified in this study; they are:

- Factor 1 (F1): Regular communication for better understanding and balance of needs between caregivers and person with cancer;
- Factor 2 (F2): Personal well-being especially on the control of emotion, communication and spiritual beliefs mainly on the quest of meaning of life and the faith in the healing process;
- Factor 3 (F3): Basic healthcare, counselling and service;
- Factor 4 (F4): Access to information related to cancer or patient care information and Services;
- Factor 5 (F5): Coping with change especially the change of life routine and perspective on life;
- Factor 6 (F6): Learning through online information and connection with others on cancer care.

These factors are aligned with the study conducted by [6] which proposed four domains of need, namely Healthcare Service (F3 of the current study), Psychological and Emotional (F2 and F5 of the current study), and Information (F4 of the current study). The last factor by [6] is Work and Social Needs. This factor was not emerged in the current study.

A. The emergence of communication as a more important need factor

The current study has revealed that Communication (F1) is a more important need factor. Communication shows the highest percentage of variance explained. Its mean value also shows the highest need compared to other factors.

Reference [5] argued that, "…effective communication for cancer patients and/or caregivers can meet information needs, reduce caregiver burden, improve physical and mental health, and promote intimacy." (p. 1) According to their critical review findings, the communication needs identified are multifaceted nature of cancer patients and/or caregivers in terms of communication target, content, style, timing, and preferences. For example, communication targets included health professionals, peers, caregivers, and patients. Communication content included illness-related, emotional support, daily life, sexuality, death, and a way to communicate with health professionals [5].

Other major contributing factors of need are Psychological and Emotional (F2) and Healthcare Service (F3). By taking emotional support as an example, evidence shows that both negative and positive emotions are important in communication [32]. A study found that cancer patients who expressed negative emotions and received an empathetic response from their oncologist perceived communication as being more favorable [31].

Reference [30] found the caregivers' needs which are: 1) accessing and understanding the information, 2) relationship with healthcare providers, 3) relationship with the care recipient, 4) managing challenges of caregiving and support systems. Other areas of need mentioned by [30] are basic healthcare, communication with care recipients, and support systems; these needs are also aligned with the findings of the current study.

According to [30], caregivers' health literacy is multidimensional. The current study has also found the need for multidimensional support for caregivers and their personal well-being. Support is not just on communication or personal well-being, but also involves other needs such as health and physical care, finance, dietary, learning and access to information and supportive people. They need guidance in communication for their preferred way or form of counselling methods. They need a credible source of information through consultation with their healthcare professionals. Caregivers need to communicate with the cancer patient and their family members.

One factor found is the need for coping with change especially the changes to the life routine and perspectives on life. The study in [27] also identified the need for coping with change. In their study, they observed that detection of "cancer creates the context for the caregiver's relationships (with patient, and the healthcare system), and the cognitive, behavioral, affective, and spiritual responses" (p.775); the caregiver's prior experiences and social support network influence the caregivers' perception towards the diagnosis, relationships, and personal responses. Mainly there is a need to cope with change in relationships among key parties who are involved in a patient-centered healthcare system, such as the healthcare providers, doctors, friends and relatives and the communities.

B. Learning need through online information and connection

The current study identified "learning through online information and social connection" as an emerging factor of need. Some input from open feedback indicated that online technologies can be used for seeking "reliable" information about cancer care, healthy lifestyle and diet. However, it is at the lower need level compared to other needs which are more critical during the caregiving period. This could be due to the support by doctors and healthcare teams to provide reliable information that are able to meet their needs.

Overall implication is that the present study has identified six-factor needs with communication need being perceived higher among all factors. The authors discovered that communication need is multifaceted, with communication target, content, style and timing should come at the right and in guided ways. Effective communication with the person with cancer, the other family members and healthcare professionals should be considered the aspect influencing caregivers' well-being. Usually, the content and style of communication are often neglected. This study found the various aspects of communication to be complex but interesting for future research.

C. Limitations

The study has sampling limitations, the caregivers are mostly from one ethnicity and the number of respondents is less than 85. Most of the samples are from the middle income group, which biases towards the needs for financial and healthcare. The convenience sample for this study comes from one hospital. The type of cancer patients and their caregivers were selected purposively from cancer clinics in a hospital, they were affected by different kinds of cancer. Hence, it is difficult to gather data of different samples based on different cancer types due to sampling limitations. All these limitations are acknowledged. Future research is proposed to incorporate more participating hospitals and caregivers.

VI. CONCLUSION AND FUTURE WORK

The study provides a systematic approach of identifying needs of cancer caregivers. The factors derived from this study have not only highlighted the common domain of needs such as Basic Healthcare, Access to Information, Psychological and Emotional (personal well-being, coping with change), but also encountered other need factors, namely communication and learning need which are corroborated with some scholar works such as [5], Yuen et al.'s works [29] [30] and the works by [14] respectively. These factors should also be considered for any policy decisions or training modules for caregivers, in line with the importance of a seamless learning model in health education. The need factors found are proposed to be incorporated into the patient-centered healthcare system for creating a more comprehensive healthcare model in cancer care involving the support towards the needs of caregivers. This study can be a source of information or as an added literature for cancer caregivers' needs.

Future directions of this research will consider to study further on one or two aspects of well-being dimensions, for example on trustable information need in decision making. Future research will also consider the different duties of caregivers and the level of quality of treatment received by patients. These variables may influence the requirement of needs by caregivers and subsequently affect the quality of life for the family.

ACKNOWLEDGEMENT

The authors would like to express gratitude for the support given to the research team by Malaysia Ministry of Higher Education (MOHE)'s Fundamental Research Grant Scheme (FRGS) support to this research work. Our appreciation also goes to National Cancer Society Malaysia (NCSM) for their support to the data collection process. Thanks to all the research respondents who were going through the experiences of giving care to their loved ones. Finally, we also would like to thank the anonymous reviewers (all of them) for providing constructive feedback to this paper.

REFERENCES

[1] M. K. Barton, "Earlier adjuvant therapy is beneficial in patients with breast and colon cancer," CA: A Cancer Journal for Clinicians 0(0): n/a-n/a. 2015. [Online]. Available from: http://doi.wiley.com/10.3322/caac.21296 (assessed 20 July 2016)

[2] S. Chen, "Information needs and information sources of family caregivers of cancer patients," Aslib Journal of Information Management, vol. 66(6), pp. 623–639, 2014.

[3] T. Carmillia and M. Sivasothy, "Quality end-of-life cancer care. Health and Living Live Forum" BFM: The Business Station. [Retrieved: August, 2020] Available from: https://www.bfm.my/health-and-living-live-2018-part-2-quality-end-of-life-cancer-care.html

[4] J. N. Dionne-Odom, A. Azuero, and K. D. Lyons, "Benefits of early versus delayed palliative care to informal family caregivers of patients with advanced cancer: Outcomes from the ENABLE III randomized controlled trial," *Journal of Clinical Oncology, vol.* 33(13), pp. 1446–1452, 2015.

[5] J. Li et al. 2020. "Communication Needs of Cancer Patients and/or Caregivers: A Critical Literature Review," Journal of Oncology, vol. 2020, 12. https://doi.org/10.1155/2020/7432849

[6] A. Girgis, S. Lambert, and C. Lecathelinais, "The supportive care needs survey for partners and caregivers of cancer survivors: Development and psychometric evaluation," Psycho-Oncology, vol. 20(4), pp. 387–393, 2011.

[7] J. F. Hair, R.E. Anderson, R. L. Tatham, and W. C. Black, "*Multivariate Data Analysis*," New Jersey: Prentice-Hall, 1998.

[8] G. F. Ho, N. A. Taib, R. K. Pritam Singh, C. H. Yip, M. M. Abdullah et al., "What if all patients with breast cancer in Malaysia have access to the best available care: How many deaths are avoidable?" Global Journal of Health Science, vol. *9*(8), 32, 2017. Available at http://doi.org/10.5539/gjhs.v9n8p32 (accessed 10 June 2016)

[9] M. Househ, E. Borycki, and A. Kushniruk, "Empowering patients through social media: The benefits and challenges," *Health Informatics Journal*, vol. 20(1), pp. 50–58, 2014.

[10] M. Kasmiah, "Creating greater awareness," *New Straits Times Life & Times*, 25 May, p. 7, 2013.

[11] E. E. Kent, "Cancer caregivers face more stress & depression," [Retrieved: June, 2018]. Available from: http://www.cancer.net/blog/2016-

09/importance-palliative-care-patients-and-caregivers-news-2016-palliative-care-oncology-symposium

[12] K. Krishna, “A heavy responsibility for Dr H. Krishna Kumar,” Fit For Life “People” *Sunday Star*, 3 August, pp.10, 2014

[13] H. W. Lee, T. Ramayah, and N. Zakaria, “External factors in hospital information system (HIS) adoption model: A case on Malaysia,” *Journal of Medical Systems, vol.* 36(4), pp. 2129–2140, 2012.

[14] C. K. Looi, P. Seow, and B. Zhang, “Leveraging mobile technology for sustainable seamless learning: A research agenda,” *British Journal of Educational Technology, vol*. 41(2), pp. 154–169, 2010.

[15] K. C. Madathil, A. J. Rivera-Rodriguez, and J. S. Greenstein, “Healthcare information on YouTube: A systematic review,” *Health informatics journal, vol.* 21(3), pp.173–94, 2015.

[16] Malaysia Ministry of Health (MOH), “National Cancer Registry Report 2007”. Putrajaya Malaysia, 2007.

[17] R. M. Razali, P. C. Bee, and G. G. Gan, “Survey of willingness to accept chemotherapy among elderly Malaysian patients,” *Asian Pacific Journal of Cancer Prevention, vol.* 14(3), pp. 2029–2032, 2013.

[18] N. Ribeiro, L. Moreira, A. M. P. Almeida and F. Santos-Silva , “Mobile Seamless Learning Tool for Cancer Education,” In: Mealha Ó., Divitini M., Rehm M. (eds) Citizen, Territory and Technologies: Smart Learning Contexts and Practices. SLERD 2017. Smart Innovation, Systems and Technologies, vol. 80. Cham: Springer, 2018.

[19] Seattle Cancer Care Alliance, “Patient & Family Education: Companion Guide for Caregivers and Family Members,” 2013. Available from: http://www.seattlecca.org/client/patient-education/SCCA-Companion-Guide-for-Caregivers.pdf (accessed 1 August 2016)

[20] P. Sercekus, D. B. Besen, and N. P. Gunusen,, “Experiences of family caregivers of cancer patients receiving chemotherapy,” *Asian Pacific journal of cancer prevention : APJCP,* vol. 15(12), pp. 5063–5069, 2014.

[21] M. Sharples and J. Roschelle “Guest Editorial Special section on Mobile and Ubiquitous Technologies for Learning,” *IEEE Transactions on Learning Technologies,* vol. 3(1), pp.4-6, 2010.

[22] A. B. Sullivan and D. Miller, “Who is Taking Care of the Caregiver?”, *Journal of Patient Experience*, vol. 2(1), pp. 7–12, 2015. Available at http://doi.org/10.1177/237437431500200103

[23] S. H. Teoh, A. C. Koo, and S. Parmjit, “Extracting factors for students’ motivation in studying mathematics,” *International Journal of Mathematical Education in Science and Technology, vol.* 41(6), 2010. 711-724. *International Journal of Mathematical Education in Science and Technology* 41(6): pp. 711–724, 2010.

[24] The Rocky Mountain University of Health Professions (RMUoHP) Biostatistics Resource,“How to use SPSS: Factor Analysis - Principal Component Analysis,” [Video file]. Available from: https://www.youtube.com/watch?v=UWP9OEoaNnE

[25] I. Veuillotte, G. Morel, and S. Pitois, “General practice and the Internet revolution. Use of an Internet social network to communicate information on prevention in France,” *Health informatics journal,* vol. 21(1), pp. 3–9, 2015.

[26] WHO Media Center, *Cancer,* 2015. Available at: http://www.who.int/mediacenter/factsheets/fs297/en/ (accessed 5 August 2016).

[27] A. Williams and M. Bakitas, “Cancer Family Caregivers: A New Direction for Interventions,” *Journal of Palliative Medicine,* vol. 15(7), pp. 775–783, 2012.

[28] M. Woodward, “A consensus plan for action to improve access to cancer care in the Association of Southeast Asian Nations (ASEAN) region,” *Asian Pacific Journal of Cancer Prevention, vol.* 15(19), pp. 8521–8526, 2014.

[29] E. Y. N. Yuen, T. Knight, L.A. Ricciardelli, and S. Burney, “Health literacy of caregivers of adult care recipients: A systematic scoping review,” *Health and Social Care in the Community*, pp. 1–16, 2016. Available from: http://doi.org/10.1111/hsc.12368 (accessed 10 August 2016)

[30] E. Y. N. Yuen, S. Dodson, and R.W. Batterham, Development of a conceptual model of cancer caregiver health literacy. European Journal of Cancer Care (December 2014), 2015. Available from: http://doi.wiley.com/10.1111/ecc.12284 (accessed 8 August 2016)

[31] C. Malhotra, R. Kanesvaran, L. Krishna, et al., “Oncologists’ responses to patient and caregiver negative emotions and patient perception of quality of communication: results from a multi-ethnic Asian setting,” *Supportive Care in Cancer,* vol. 26(3), pp. 957–965, 2018

[32] A. L. Terrill, L. Ellington, K. K. John, et al., “Positive emotion communication: fostering well-being at end of life,” *Patient Education and Counseling*, vol. 101(4), pp. 631–638, 2018.

[33] Cancer.Net Website [Retrieved: August, 2020] https://www.cancer.net/ cancer.net

[34] Skill For Care Website [Retrieved: August, 2020] https://www.skillsforcare.org.uk/About/About-us.aspx

[35] Dignity In Care Website. [Retrieved: August, 2020] https://www.dignityincare.org.uk/About/

TABLE 2. AN ANALYSIS OF THE TOTAL VARIANCE EXPLAINED BY COMPONENTS.

Total Variance Explained									
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	16.714	43.984	43.984	16.714	43.984	43.984	5.537	14.571	14.571
2	2.483	6.536	50.520	2.483	6.536	50.520	5.267	13.862	28.433
3	2.211	5.817	56.337	2.211	5.817	56.337	4.016	10.568	39.001
4	1.504	3.957	60.294	1.504	3.957	60.294	3.408	8.969	47.970
5	1.270	3.342	63.636	1.270	3.342	63.636	3.382	8.900	56.870
6	1.187	3.124	66.760	1.187	3.124	66.760	2.738	7.206	64.076
7	1.121	2.949	69.709	1.121	2.949	69.709	1.664	4.379	68.455
8	1.031	2.712	72.421	1.031	2.712	72.421	1.507	3.967	72.421
9	.938	2.468	74.889						
10 **	.924	2.431	77.320**						
Extraction Method: Principal Component Analysis.									

** This table only shows up to 10 components, and did not show all 38 components or up to 100% cumulative value of total variance explained due the space limitation. The bold numbers indicate an acceptable value for the percentage of variance explained for this study. The items cluster into these six groups defined by the highest loading on each item.

TABLE 4. FACTORS WITH ITEMS LOADED TO THE SIX-FACTOR OF NEEDS.

Factor # and its interpretations (Factor Name)	Item # and Item Statement	Mean (Stad. Deviation) Internal reliability Cronbach's Alpha, α
F1) Communication and Balance of Needs (Exchange of ideas / thoughts) Total items loaded = 6	13. Maintain regular communication with your family	2.76 (.86137) Mean near to 3, moderate need α = .896
	14. Need of regular communication with the person with cancer	
	19. To understand the experience/situation of the person with cancer	
	20. Balance your needs with the needs of the person with cancer	
	29. Having opportunities to participate in decision making on the treatment plans with cancer patient	
	15. Communicating to other caregiver(s)	
F2) Personal well- being (Spiritual, health, and emotion) Total Items loaded = 6	10. Look after your own health, including eating and sleeping properly	2.59 (.79164) α =.902
	25. Coping with the person with cancer's recovery not turning out the way you expected	
	27. Finding your spiritual beliefs	
	28. Finding the life meaning during the process of caregiving	
	30. The needs of belief/faith in the healing process	
	36. Learn how to cope with your emotion	
F3) Basic Healthcare and service (Services such as medical, counselling, or community support) Total items loaded = 7	9. Seek help in managing stress in the person with cancer	2.65 (.80909) α = .915
	6. Be involved in the person with cancer's care, together with the medical team	
	7. Have opportunities to discuss your concern with the healthcare providers on cancer	
	12. Get counselling service from the professional counsellor	
	5. Access community healthcare services when needed	
	8. Make sure complaints from the person with cancer's care are properly addressed	
	11. Adapt to the change of routine activities to cater to the needs of the patient	
F4) Access to information related to cancer and cancer care Total items loaded = 4	1. Accessing information relevant to your need as a caregiver	2.60 (.80015) α = .885
	2. Accessing information about the person with cancer's prognosis, or likely outcome (the outcomes of getting the illness)	
	3. Access information about support services for cancer caregivers	
	4. Access information on what the person with cancer's physical/emotional needs	
F5) Coping with change and fear of change Total items loaded = 3	23. Working through your feelings about death and dying	2.51 (.80573) α = .786
	24. Coping with others who do not acknowledge the change on your life as a result of caregiving	
	17. Address the feeling of fear of cancer may occur to other family member	
F6) Learning through online information and connection Total items = 3	34. Learn from online healthcare materials in making decision(s) for the person with cancer	2.26 (.79502) Lower mean, near to scale 2, which means lower need α = .822
	37. Seek information through social media site (eg Facebook, Twitter, Blog) about healthcare	
	38. Connecting with people via internet to learn about healthcare matters	

Notes: 1 denotes no need (either satisfy or non-applicable), 2 denotes low need, 3 denotes moderate need, and 4 denotes high need.

Copyright (c) IARIA, 2020. ISBN: 978-1-61208-817-4

Lessons Learned from the Development of Mobile Applications for Fall Detection

Ítalo Linhares, Rossana M. C. Andrade, Evilasio C. Júnior, Pedro Almir Oliveira,
Breno Oliveira and Paulo Aguilar

Group of Computer Networks,
Software Engineering and Systems (GREat)
Federal University of Ceará
Fortaleza, Ceará, Brazil
Email: {italoaraujo, evilasiojunior, pedromartins,
brenooliveira, pauloaguilar}@great.ufc.br, rossana@ufc.br

***Abstract*—Falls are the leading cause of older adults injuries and solutions are needed to address this issue. One way to meet this is by developing applications that use sensors embedded in devices like smartphones and smartwatches. This paper presents our experience in developing such applications and the lessons learned during their development and evolution. First, we developed an application called fAlert to identify a fall using data from a smartphone's accelerometer. However, the usage of this kind of mobile device for detecting falls is not natural, because it needs to be positioned at the level of the user's chest. Then, we developed a new app called WatchAlert, which runs in smartwatches. In that case, we also created an algorithm that uses two sensors, accelerometer and gyroscope, and later evolved it to use only the accelerometer with better results. Moreover, we use first a fall detection threshold algorithm in this solution. Next, we expanded this strategy to use threshold and machine learning algorithms, which were evaluated considering the accuracy, false negative, and time criteria as well as their features. We believe that this study can support the development of new systems and devices for detecting falls. As future work, it would be interesting to assess the related energy cost of the fall detection approaches studied.**

Keywords–Falls; Lessons; IoT Health; IS for healthcare.

I. Introduction

According to the World Health Organization (WHO), falls are the leading cause of injuries in older adults. About 28-35% of the people aged 65 years and over fall every year, and this number increases when the person's age is over 70 years old, achieving 32-42% [1]. Also, falls cause 40% of all injury deaths [1]. These facts represent alarming data, and an effort is needed to decrease this number. One way to meet this need is preventing or quickly identifying older adults falls.

There are many ways to detect falls, such as with sensors or computer vision. In case of sensors, they can be embedded in smartphones, smartwatches, and other devices like wearables to monitor the vital signs and the activities performed by the user [2]–[5]. Considering computer vision, we can use image processing to identify if a person has fallen or if he/she is performing another activity [6].

Although we can use computer vision, many users may feel uncomfortable because their activities can be recorded to be processed. Still, users may not trust data storage and the permissions of those who will access the images. Then, we perceive an impact on the usage of these systems related to reliability issues, and we should use a way less invasive to users, which can be using the sensors previously cited.

Based on this, we can find in the literature efforts to detect a fall using sensors embedded in wearables devices such as smartphones and smartwatches as in [2] [7]–[10] or mixing data of wearables with other devices as in [11]. Examples of sensors present in these devices that can help detect a fall are the accelerometer and the gyroscope to check the user movement, and the magnetometer that can aid in the identification of the fall location.

Thus, we developed two applications and three thresholds-based algorithms to detect a fall. The first solution developed is the fAlert application [4], which runs in smartphones and uses accelerometer and magnetometer sensors. As the device should be situated in the user's chest, its usage is not natural to the older adult. So, we sought to evolve the solution using smartwatches.

This solution is the WatchAlert application [5], which, in the first version, was developed with two sensors, accelerometer and gyroscope, to detect a fall. We continued looking for improvements. Then, we observed that we could evolve the algorithm to use only one sensor without causing a significant reduction in its accuracy, which indicates how capable the algorithm is to identify a real fall.

Considering our experience in the development of such fall detection solutions, including the use of thresholds and machine learning techniques, and service-oriented architecture, we present the evolution of these algorithms, along with the motivations to do this and the lessons learned in this process. We consider as the main contribution of this paper is to compile and discuss the top six lessons learned while developing fall detection solutions to facilitate future research. These lessons could guide the developers of the detection fall solutions to avoid the same problems that we faced. Then, this contribution can improve the Software Engineer area to support the e-health solutions.

This paper is divided as follows: Section II discusses related work. Section III shows the fAlert and WatchAlert applications in addition to the fall detection algorithm versions. Lessons learned are discussed in Section IV. Finally, Section V brings conclusions and future directions.

II. Related Work

As mentioned in the previous section, there are several works focused on fall detection [12]. The significant public health problem generated by falls, especially in the elderly population, can justify this large number of papers. However, it is tough to find papers with lessons learned about the

Copyright (c) IARIA, 2020. **ISBN: 978-1-61208-817-4**

development of fall detection solutions. In general, each paper focuses on presenting, in a particular way, its fall detection approach and the results obtained with this model. Moreover, the absence of standardization regarding sensors, algorithms, and evaluation methods makes a comparison between these works a hard task.

In our research, we find just a recent paper with lessons learned related to fall detection systems [13], but there are other works presenting fall detection applications similar to those created throughout this research [14]–[17].

The work in [13] presents challenges concerning the creation of a multimodal database for fall detection data. The data acquisition system created obtains data from infrared sensors, wearable sensors, cameras, and an ECG (electrocardiogram) monitor. The authors state that the lessons learned are fundamental to database consolidation.

Different approaches use pre-processing data from the accelerometer to select the best features for the fall detection systems [15] [17]. These studies work with threshold-based algorithms, but also use other strategies to achieve better results. In [15], the proposed system, besides the accelerometer, utilize gyroscope and magnetometer in a secondary threshold verification to achieve more accuracy. [17] uses optimization algorithms (genetic algorithms, and simulated annealing) to set thresholds.

Some researches advocate the use of machine learning algorithms to detect falls. The system proposed in [14] presents a fall detection solution that uses an accelerometer and Supported Vector Machines (SVM). The authors propose a data pre-processing, combining features to find a better combination for the SVM method. Moreover, the study confronts the machine learning algorithm proposed with a threshold-based solution, and, according to the authors, when the parameters of falling and non-falling are very close, the results of the SVM method are better than the threshold-based algorithm.

However, instead of using either approach, some authors prefer to combine both thresholds and machine learning algorithms in a fall detection solution. The work in [16] proposes a design model to create a healthcare monitoring system based on wearables. This model combines both threshold and machine learning algorithms (decision tree, SVM, and k-Nearest Neighbor) to detect a fall.

The study in [13] addresses the difficulties related to building a fall database, which is important for fall detection applications but does not address the process of developing these applications. The papers [14]–[17] present different approaches to create fall detection systems. Each approach contributes to understanding the challenges and possible solutions in the design and modeling processes of these systems. However, there are other challenges during the development process that the researches do not show.

In contrast, we believe that our work stands out for presenting lessons learned, focusing on good practices for the development of fall detection applications. In this article, we discuss issues related to the design, modeling, and development processes. Besides, we expose lessons learned for fall detection strategies based on thresholds and in machine-learning techniques.

These lessons learned come from the experience of several years in the study and development of fall detection applications. They can assist the process of developing new solutions and patterns by indicating important steps to be followed and difficulties common to the process of creating this type of application.

III. Evolution of Applications

A. *fAlert*

We show the first solution that we developed to detect a fall. As said in the previous section, we propose an algorithm to analyze the user acceleration with smartphone support. The behavior was analyzed with an algorithm whose input is the data collected from the accelerometer and the magnetometer.

As we saw previously, the accelerometer data allows the calculation of acceleration and the comparison with the acceleration of gravity. The collision with the floor can also be calculated with the acceleration, but for the verification to be correct in the fAlert algorithm, the time window is 1.5 seconds.

In sequence, the algorithm checks if the device is in a position of 45 degrees about the floor. If the answer to this verification is positive, the algorithm detects a fall, and it asks the user if he/she is well. He/she has 25 seconds to answer, and in negative case, fAlert sends a message to a caregiver.

We evaluated this algorithm considering only one daily activity, the user laying, and some fall scenarios. We executed the experiment 30 times with four combinations of the thresholds, changing the value of detection of a collision with the floor. As values for the thresholds, we have the threshold for free fall equal to 0.2 times of the gravity acceleration (0.2G), and the threshold to determine the collision equal to (1.5G). This setting resulted in 93% of detected falls, and 67% of the daily activities recognized.

B. *WatchAlert*

We developed the WatchAlert application to run in a smartwatch, collecting data from the accelerometer and the gyroscope in the first version and using only the accelerometer in the second version.

1) Architectural Aspects: Although we know the device constraints, we do not know what the impact in the application is. When we started the development of WatchAlert, we decided that the algorithm (collecting and processing) should run entirely in the smartwatches. When we started the software testing phase of this first version of the algorithm, the application would break when it collected and processed the data simultaneously.

A solution adopted by us was the code division to run the collect task in the smartwatch and processing in the smartphone connected with the smartwatch. With this division, in the flow shown in Figure 1, the only activity performed in the smartwatch is “Collect the data”.

With this problem solved, we can perform the tests in the WatchAlert application, and we also verify that the collection was not continuous, and it does not allow the need to define services to run in the smartwatch to allow the collect.

2) First Version of the Algorithm: We defined an algorithm based on [2], and we used the data from the accelerometer and gyroscope, combining them to detect a fall. To do this, the steps performed in the algorithm verify initially if the user suffers a free fall from the analysis of the three axes of the accelerometer according to (1). If the values obtained with the calculation is lower than 5,3936, the first signal of a fall is detected.

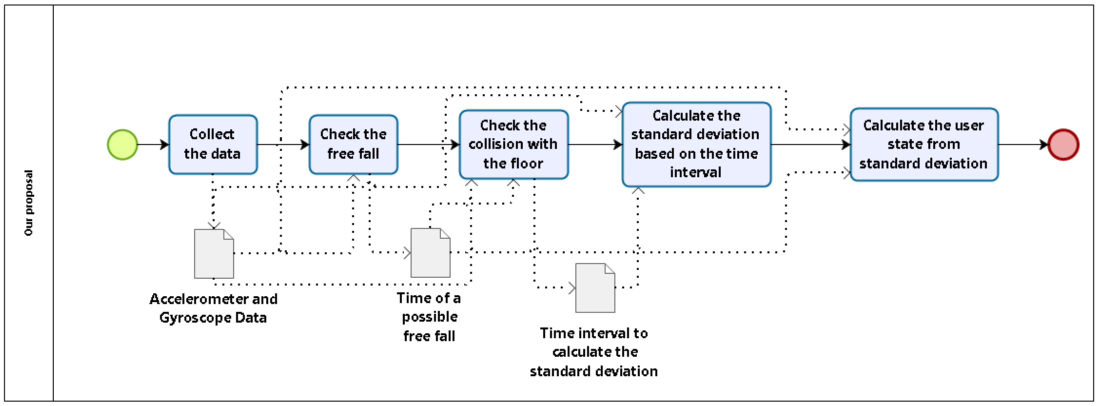

Figure 1. Flow of the second version of WatchAlert algorithm.

$$VA(t_i) = \sqrt{{A_x}^2(t_i) + {A_y}^2(t_i) + {A_z}^2(t_i)} \tag{1}$$

In sequence, the algorithm checks the user impact on the floor looking for another evidence, also using (1), and the result should be higher than 23,5359. This procedure is verified considering the interval of 0.4 seconds after the free fall. The third step of the algorithm, different from the previous steps, uses the data from the gyroscope to determine the movement performed by the user's arm. The gyroscope data contains three axes like the accelerometer, but the difference is that it calculates the angle of rotation. If the user moves the arm faster than 250°/s in the interval between the free fall and the impact, we have another sign of a fall. Equation (2) allows us to know the angle of movement.

$$GS(t_i) = \sqrt{{G_y}^2(t_i) + {G_z}^2(t_i)} \tag{2}$$

Following the steps, we return to use the accelerometer data to continue finding fall evidence. At this moment, the algorithm calculates the standard deviation from the free fall until the collision with the floor to verify if the resultant value is similar to the standard deviation of a fall previously defined.

The last step that allows us to diagnose a fall is the user's state after a possible fall. We should consider three scenarios: i) user is active; ii) user is inactive; and iii) user is partially active. In the first situation, we have the user performing a movement as clapping or jumping, and the accelerometer behavior is like a new impact. If the user is inactive, he/she falls and is immobile, needing fast support. Moreover, the last one, the user maintains a movement, but differently from the active state, he/she makes less abrupt movements.

To calculate this, we compare the result of the sum of the acceleration determined by (3) for 1.5 seconds, starting in the moment of the impact verified in the second step. If the result is greater than 1274,86, the algorithm identifies a fall.

$$SA_i = \sum_{j=1}^{N} [|A_x(j)| + |A_y(j)| + |A_z(j)|] \tag{3}$$

3) Second Version of the Algorithm: We observed the algorithm described in Section III-B2 and performed an improvement to use only the accelerometer data from smartwatches. In comparison with the previous version of the algorithm, we made two main changes: i) remove the gyroscope step; and ii) changes in the equation of the last step to identify the fall.

The first change is motivated by the evolution of our algorithm. About the second main difference, we performed it based on the analysis during tests of the first version because raw data can be too imprecise. To improve this, we use the standard deviation considering the same time interval, 1.5 seconds after the possible collision with the floor.

As a consequence of all these improvements, we changed thresholds used to determine a fall. We performed an experiment to detect these new values with data collected (data available in: https://bit.ly/31wNiiQ) using the LG Urbane smartwatch and performed with five persons, whose profile is described in Table I, and we found: i) in the first step, was 6.864655. ; ii) in the second step, was 29.41995; iii) in the third step, was 9.80665; and iv) we change to 4.903325 in the fourth step.

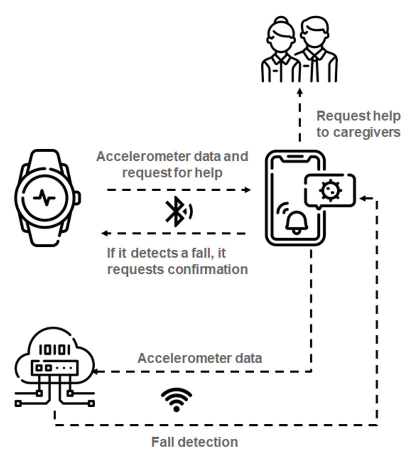

Figure 2. Overview of the third version of WatchAlert.

C. Machine Learning Algorithm

Both algorithms previously presented use thresholds-based fall identification strategy. This strategy verifies the combination of a set of measurements that, when exceeding certain thresholds, can characterize a possible fall event. Another

TABLE I. PROFILE OF PARTICIPANTS DURING DATA COLLECTION.

	P1	P2	P3	P4	P5
Gender	M	M	M	M	F
Age	28	30	23	20	20
Height	1.78m	1.72m	1.73m	1.83m	1.64m
Weight	70Kg	90kg	90Kg	82Kg	55Kg

strategy that can be used to detect falls is the usage of supervised machine learning algorithms. We have used this new approach to create the third version of WatchAlert. Figure 2 shows an overview of this version.

We evaluate the machine learning algorithms and obtained more accuracy than threshold algorithms. However, it is important to highlight that the usage of machine learning requires a higher processing power than the strategy with thresholds. This processing power can be unfeasible for some smartphones, and we used the cloud to execute this algorithm. Considering the message exchange, total data processing time, and latency, we achieved a low response time to trigger the alarm, using a simple WiFi internet connection and 4G. We believe that response time is within an acceptable margin. Below, we detail the selection and evaluation of the features and the machine learning algorithm.

It is worth mentioning that in this third version of WatchAlert, if the application is offline, it not stop working. In this case, WatchAlert behaves as in its second version, using the threshold approach that is less expensive for the mobile device and also allows to identify possible falls with slightly less accuracy than the Machine Learning approach.

1) Feature Selection: One of the first activities to create a new machine learning model is to select the features that can be used to characterize the relevant events. These features are values based on treatments or calculations made on the raw data, which we obtained from the smartwatch accelerometer.

TABLE II. FEATURES EQUATIONS.

Measure	Equation
Mean	$(V_1 + V_2 + V_3 + ...V_{n-1} + V_n)/n$
Minimum (MIN)	$Min(V_1, V_2, V_3, ...V_{n-1}, V_n)$
Maximum (MAX)	$Max(V_1, V_2, V_3, ...V_{n-1}, V_n)$
Standard Deviation (STD)	$\sqrt{\frac{\sum_{i=1}^{n}(V_i - Mean)^2}{n}}$
RMS	$\sqrt{\frac{1}{n}\left(V_1^2 + V_2^2 + V_3^2 + ...V_{n-1}^2 + V_n^2\right)}$
Kurtosis	$\frac{\sum_{i=1}^{n}(V_1 - Mean)^4}{\left(\sum_{i=1}^{n}(V_1 - Mean)^2\right)^2} - 3$

As shown before, the (1) is used to combine the three axes values of the accelerometer. Based on resulting value $VA(t_i)$ for each measurements set during a given period, we calculate the following features: mean, maximum, minimum, standard deviation, RMS, and Kurtosis. Table II presents the equations referring to the calculation each one of these features. In this table, V_i is a value of (1) for measurement i, and n is the measurement number.

These six features are used to represent the classes of events that we want to detect (fall and no-fall). The choice of these features was made based on the literature, since they are widely used by several other surveys of fall detection solutions, such as [18]. We emphasize that the evaluation process of these features is essential to choose only the features that collaborate with the result. Below, we present the machine learning algorithms examined and the feature analysis.

2) Algorithms and Feature Evaluation: First, we performed a correlation analysis among features to remove those that were directly correlated. We understand that this correlation depends on the data, but keep this behaviour among the features can reduce the performance of the machine learning classifiers. Thus, first it was applied the Shapiro-Wilk normality test [19]. At the 0.05 significance level, the data was not significantly drawn from a normally distributed population. Because our data have just ordinal fields and it is not normally distributed, we decided to use the Spearman correlation [20]. Table III presents the results obtained using the OriginPro Data Analysis and Graphing Software v9.1 [21]. Considering these results, we noticed a high correction between RMS and STD (0.76955), RMS and MEAN (0.88088), RMS, and MAX (0.70221), MAX and STD (0.87859). Hence, we decided to remove the Standard Deviation (STD) and Mean (MEAN) features of this evaluation, keeping the Maximum (MAX) because it is very relevant to the previously presented threshold algorithm.

TABLE III. SPEARMAN CORRELATION.

	STD	MEAN	MAX	MIN	RMS	Kurtosis
STD	1	0.4575	0.87859	-0.74941	0.76955	-0.31087
MEAN	0.4575	1	0.4587	-0.00775	0.88088	-0.33049
MAX	0.87859	0.4587	1	-0.63796	0.70221	0.0596
MIN	-0.74941	-0.00775	-0.63796	1	-0.3261	0.07504
RMS	0.76955	0.88088	0.70221	-0.3261	1	-0.40499
Kurtosis	-0.31087	-0.33049	0.0596	0.07504	-0.40499	1

After feature selection, it was possible to evaluate the performance of two algorithms: Support Vector Machines (SVM) [22], and Random Forest [23]. We selected these algorithms because some previous tests showed that they present better results when compared with other algorithms such as MLP, J48, Random Tree, and Naive Bayes. Furthermore, as many works in the literature have already analyzed many classifiers for the fall detection problem [24], we have decided to deepen this analysis considering only these two algorithms. Regarding SVM, we used three different kernels: Linear, Polynomial, and Gaussian.

These algorithms - Random Forest and SVM - were executed considering classifiers with attribute selection and with attribute selection combined to cost-sensitive learning [25]. For the Random Forest attribute selection, we used Information Gain. SVM attribute selection was performed with SVM Attribute Evaluation [26]. The cost matrix $[[0, 5], [1, 0]]$

used in the Cost-Sensitive Classifiers penalized five times more the false negatives, *i.e.*, cases in which a fall occurred and the classifier labeled no-fall. Considering the negative impact caused by falls, failure to detect a fall is extremely dangerous.

Regarding datasets, we used an original dataset with 359 no-fall instances and 67 fall instances. We understand that this number of instances represents a small dataset, but it is important to highlight the difficulty of getting large and real data related to elderly falls. Thus, we argue that this dataset can provide evidence about the feasibility of our approach. So, considering feature correction analysis, we have created a similar dataset, but without standard deviation (STD) and mean (MEAN) features. Finally, due to the class unbalancing, we have created a third version of the dataset applying the Synthetic Minority Oversampling TEchnique (SMOTE) [27]. This last dataset has 359 no-fall instances and 268 fall instances. All experiments were performed on the WEKA data mining software tool [28], and each classifier was executed 30 times.

Table IV presents the results considering accuracy, precision, recall, true positive rate (TPR), false-negative rate (FNR), false-positive rate (FPR), true-negative rate (TNR), elapsed time training (in seconds), and elapsed time testing (in seconds). We were seeking for a classifier with high accuracy, high recall, and low false-positive rate, because it is crucial to classify data correctly (effect analyzed by accuracy), avoiding the non-identification of a true fall (effect analyzed by recall), and reducing the number of false fall events alarms (effect analyzed by false-positive rate). The Equations (4), (5) and (6) present how these metrics are calculated [29].

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

The best result was obtained using a Random Forest algorithm with Information Gain attribute selection, and considering the original dataset with SMOTE. We got 95.13% of accuracy, 95.13% of recall and 5.20% of false-positive rate. With these results, it is important to highlight that still remains a gap for future studies regarding the improvement of accuracy and recall, and the reduction of false-positive rate. In a real context, with a lot of no-fall events, 5.20% of fall alerts in regular situations can reduce users' adherence to this kind of monitoring. As in this work, we used only the default algorithm parameters, these results can be improved considering a fine-tuning of the parameters.

To conclude our analysis, we also checked the average time spent by each classifier in training and testing. We observed that the Random Forest time for both training and testing activities was not the best. However, the values did not differ much from the best results. In this case, the Random Forest with Information Gain Attribute Selection takes 0.17 seconds for training and 0.00175 seconds for testing. Because of this evaluation, we decided to use a Random Forest classifier in the WatchAlert third version.

IV. Lessons Learned

We describe in this section the lessons learned during the development of the applications fAlert and Watchalert in three versions of algorithms and the features used in them. During the development of the applications, we tested the algorithms in two ways: i) offline, which is when we implement the algorithm in a device not used to collect and detect a fall; ii) online, which is the use of the algorithm in the real device to monitor the user.

The first way, offline mode, is adequate when we want to know the number of falls and daily activities correctly identified based on the data that we collected using smartwatches. However, we can know this information, but we need to know how the algorithm runs in a real device and the consequences of its use. Furthermore, with our data analysis and considering both ways, we learned six lessons described hereafter. These lessons are described with a generic problem that we face, followed by the experienced situation and the lesson learned during our work.

LL01: Use Flow-based Programming

- **Problem:** Considering the sending of batch data, one problem occurs when the free-fall occurs in a batch, and the impact with the floor and the user state are in another. Then, the fall detection behavior can be compromised and the emergence service or caregivers cannot be contacted. How to deal with this situation?
- **Experience:** The problem was experienced during the development of the first version of WatchAlert. The accelerometer data from smartwatch was collected and sent in batch to smartphone, and several times the fall detection results was a daily activity, generating the false-negative response. In our studies to detect the cause of the problem, we identified that the batch data compromised the results, because the free fall was in different batch of the floor impact or the user state. Then, we search for the best way to correct the problem, and the solution found was the Flow-based Programming.
- **Lesson:** Considering our experience, we suggest the developers of the fall detection solutions use flow-based programming because it allows the sequenced data to avoid the improvement of the false-negative rate. This strategy is especially important in these situations because the fall detection solution deal with real-time data and needs of the precise and correct result to avoid the severe sequelae with the user.

LL02: Use Service-Oriented Architecture (SOA)

- **Problem:** As each device perform several actions, for instance, the smartwatch collects the data and interact with the user to receive the answer of the real status or the smartphone, which gets the data, process it and sends the message to caregivers or emergence service, it is important to deal with this. How can the developers of the fall detection solutions maintain several functionalities without one impact in another?
- **Experience:** In the first version of WatchAlert, we tried to run the code in sequence, according to the fall detection steps. However, there is a problem when the smartwatch continues collecting the data and wait for the smartphone response; then, the fall detection

TABLE IV. TABLE WITH ACCURACY AND FALSE-NEGATIVE RATE OF EVALUATED CLASSIFIERS.

Dataset	Classifier	Classifier Detail	Accuracy	Precision	Recall	TPR	FNR	FPR	TNR	Time Training (s)	Time Testing (s)
Original	Attribute Selected Classifier	RandomForest (InfoGain)	93.85	93.94	93.85	93.85	6.15	25.09	74.91	0.11	0.00093
		SVM (SVMAttributeEval) - Linear Kernel	93.46	93.70	93.46	93.46	6.54	20.24	79.76	0.20	0.00076
		SVM (SVMAttributeEval) - RBF Kernel	91.27	91.21	91.27	91.27	8.73	41.42	58.58	0.06	0.00155
		SVM (SVMAttributeEval) - Polynomial Kernel	87.21	90.33	87.21	87.21	12.79	23.06	76.94	53.49	0.00064
	Cost Sensitive Classifier	AttributeSelected (InfoGain): RandomForest	93.43	93.73	93.43	93.43	6.57	19.09	80.91	0.10	0.00078
		AttributeSelected (SVMAttributeEval): SVM - Linear Kernel	88.83	91.34	88.83	88.83	11.17	15.76	84.24	0.41	0.00063
		SVM - Linear Kernel	88.64	91.28	88.64	88.64	11.36	15.65	84.35	0.43	0.00203
		SVM - Polynomial Kernel	83.10	88.71	83.10	83.10	16.90	20.75	79.25	53.59	0.00141
		SVM - RBF Kernel	87.77	88.58	87.77	87.77	12.23	30.90	69.10	0.02	0.00173
Original - (STD, MEAN)	Attribute Selected Classifier	RandomForest (InfoGain)	93.62	93.69	93.62	93.62	6.38	25.30	74.70	0.10	0.00093
		SVM (SVMAttributeEval) - Linear Kernel	93.60	93.86	93.60	93.60	6.40	19.96	80.04	0.21	0.00096
		SVM (SVMAttributeEval) - RBF Kernel	89.86	90.05	89.86	89.86	10.14	49.71	50.29	0.05	0.00184
		SVM (SVMAttributeEval) - Polynomial Kernel	80.23	86.47	80.23	80.23	19.77	30.09	69.91	53.51	0.00047
	Cost Sensitive Classifier	AttributeSelected (InfoGain): RandomForest	92.49	93.05	92.49	92.49	7.51	19.03	80.97	0.10	0.00047
		AttributeSelected (SVMAttributeEval): SVM - Linear Kernel	89.06	91.55	89.06	89.06	10.94	15.57	84.43	0.34	0.00143
		SVM - Linear Kernel	88.92	91.47	88.92	88.92	11.08	15.48	84.52	0.42	0.00126
		SVM - Polynomial Kernel	64.50	82.20	64.50	64.50	35.50	29.73	70.27	53.43	0.00124
		SVM - RBF Kernel	87.28	87.51	87.28	87.28	12.72	36.57	63.43	0.02	0.00156
Original + SMOTE	Attribute Selected Classifier	RandomForest (InfoGain)	95.13	95.27	95.13	95.13	4.87	5.20	94.80	0.17	0.00175
		SVM (SVMAttributeEval) - Linear Kernel	88.47	88.72	88.47	88.47	11.53	12.35	87.65	0.31	0.00110
		SVM (SVMAttributeEval) - RBF Kernel	92.09	92.33	92.09	92.09	7.91	8.19	91.81	0.06	0.00277
		SVM (SVMAttributeEval) - Polynomial Kernel	88.49	88.98	88.49	88.49	11.51	11.81	88.19	36.29	0.00189
	Cost Sensitive Classifier	AttributeSelected (InfoGain): RandomForest	93.11	93.65	93.11	93.11	6.89	5.97	94.03	0.15	0.00064
		AttributeSelected (SVMAttributeEval): SVM - Linear Kernel	76.48	81.66	76.48	76.48	23.52	19.45	80.55	0.33	0.00139
		SVM - Linear Kernel	76.40	81.54	76.40	76.40	23.60	19.56	80.44	0.29	0.00295
		SVM - Polynomial Kernel	87.52	88.60	87.52	87.52	12.48	11.45	88.55	35.51	0.00109
		SVM - RBF Kernel	89.81	91.26	89.81	89.81	10.19	8.43	91.57	0.03	0.00281

was impacted, leading to a false-negative situation. To solve this, we tried to put services in the smartwatch, even with hardware constraints, and it works. The services were to collect data, communicate with the user, and connect with the smartphone. We perform the same with a smartphone to avoid any problem like a smartwatch, and the WatchAlert works well.

- **Lesson:** The lesson learned was that we should develop an application using Service-Oriented Architecture [30], dividing each functionality in service. For instance, we can run the data collection and check the user‘s health in the smartwatch. This avoids the increase of the false-negative rates, turning the application more accurate.

LL03: Divide the code in different devices

- **Problem:** Wearables devices have constraint hardware, which allows the use of a few services and algorithms. Then, these devices cannot collect and process the data and call the user’s caregivers. How we deal with this situation without to impact on the result?
- **Experience:** We note a problem with functionalities and the accuracy of the algorithms when we use only the smartwatch. As we deal with an IoT solution, we consider dividing the algorithms in different devices of the way that the smartwatch collects the data, sends them to the smartphone, which goal is to process it, and returns the result to the smartwatch. Furthermore, we also divide the code between the smartphone and Cloud to improve the detection and use of Machine Learning algorithms. Then, the problem was solved, and we do not suffer any impact with the time of fall detection.
- **Lesson:** The lesson learned with the experienced situation was the division of the code between different devices according to each hardware configuration. This strategy is possible due to the IoT solutions, which are composed of several devices and sensors with different configurations. Then, with this lesson, the developers of fall detection applications can better use the devices and do not suffer any impact on the accuracy of the algorithm neither in any other important service of the application.

LL04: Use offline methods of fall detection integrated with ML algorithms

- **Problem:** Health applications are critical, and the information should always be available. However, if the application uses only algorithms in the cloud, the solution can have a problem when the connectivity with the Internet and the service is inaccessible. Then, how can we assure the high availability of the service and a self-adaptive detection model?
- **Experience:** We tried to use a cloud-based solution to execute our ML approach, but this strategy is affected by network problems. Considering the criticality of this type of application, it is also essential to use an offline method to ensure service availability even with network problems. This way, we can guarantee a seamless service that evolves using new data even without internet connectivity. As smartwatches are frequently associated with a smartphone, the connection is Bluetooth, and it continues even with a problem with an Internet connection, which allows us to maintain the service.

- **Lesson:** If you want to use ML algorithms for fall detection, it should be in a cloud, and it is crucial to have an alternative method, until threshold algorithm, that does not depend on Internet access. Thus, we suggest putting at least another algorithm in the smartphone to run in the situations of Internet problems. Therefore, it is possible to ensure high service availability even that the application lost the internet connection.

LL05: Analysis and Selection of the best features for the ML algorithm

- **Problem:** When using machine learning for fall detection, we can use several features and not all features contribute to improve the result. Moreover, the greater the number of features may consequently make the algorithm processing slower. Then, how can we select better features for the machine learning algorithm?
- **Experience:** We use a set of features as an input of ML algorithms as root mean square, mean, max, and min values. However, we note that the results could be improved. In the first step to detect the best improvement, we analyzed the correlation between features used, which showed us the need to change the features selected and demonstrated in the Section about ML algorithms. Consequently, when we execute the ML algorithms, the results were improved, and the processing time was shorter.
- **Lesson:** For best results with the ML algorithms, the features used must be selected using specific techniques for this analysis considering the application scenario. We recommend using correlation analysis and attribute analysis algorithms, such as Information Gain and SVM Attribute Evaluation. In the fall detection scenario, we suggest using the maximum, minimum values, RMS, and kurtosis, when use only accelerometer data.

LL06: Choose the most suitable ML algorithm

- **Problem:** Many times, we see works that use ML algorithms for fall detection but do not explain the chosen. However, some factors impact the algorithm results, e.g., the feature type, the amount of data, and how this data is organized. Then, what algorithm should we choose?
- **Experience:** Firstly, we choose the ML algorithms for fall detection apps based on literature, but we decide to test other ML algorithms when we change the features. To our surprise, the results were different from the literature knowledge and the tree-based algorithms, like Random Forest, which is not most used as the best algorithm for detection fall scenario considering the accelerometer data that we have. Then, for each situation, several algorithms should be tested.
- **Lesson:** Choose the ML algorithm for your fall detection applications based on the data available. In the fall detection scenario, we suggest at least evaluate variations of SVM algorithm and tree-based algorithms when work with only the accelerometer data.

V. CONCLUSION

This paper discusses issues related to the development of mobile fall detection applications. We presented two applications and three algorithms developed by the GREat research group. Throughout this process, we have built up a good experience on this topic. Thus, it was possible to describe the six main lessons learned about fall detection by smartphones and smartwatches.

In short, we observed that the usage of flow-based programming and Service-Oriented Architecture (SOA) could assist in the development of fall detection applications due to a significant amount of data that needs to be processed and to keep the data capture service independent of handling them. Also, due to the smartwatches hardware restrictions, it is crucial to delegate to the smartphone the processing of data. Finally, for cases that involve cloud-based machine learning approaches, it is essential to have an alternative method that does not depend on Internet access.

In the scenario of fall detection using accelerometer data, we observe that maximum and minimum values, RMS and kurtosis are the most important features. In sequence, on the contrary to the literature, we identified that the Random Forest is better than SVM considering the machine learning algorithms found in the literature.

Thus, we believe that by following these guidelines, it is possible to create a robust, seamless, and easy-to-maintain fall detection service using sensors present on smartphones or smartwatches. As future work, we intend to evaluate the battery and latency, and comparing the thresholds and the machine learning algorithms.

ACKNOWLEDGMENT

The authors would like to thank CNPQ for the Productivity Scholarship of Rossana Maria de Castro Andrade DT-2 (N^o 315543 / 2018-3) and Coordination of Improvement of Higher Level Personnel - Brazil (CAPES) that provided to the Evilasio Costa Junior a Ph.D. scholarship.

REFERENCES

[1] World Health Organization, “Global report on falls prevention in older age,” https://www.who.int/ageing/publications/Falls_prevention7March.pdf, 2007, [retrieved: October, 2017].

[2] S.-L. Hsieh, C.-C. Chen, S.-H. Wu, and T.-W. Yue, “A wrist-worn fall detection system using accelerometers and gyroscopes,” in Networking, Sensing and Control (ICNSC), 2014 IEEE 11th International Conference on. IEEE, 2014, pp. 518–523.

[3] P. Kostopoulos, T. Nunes, K. Salvi, M. Deriaz, and J. Torrent, “F2d: A fall detection system tested with real data from daily life of elderly people,” in 2015 17th International Conference on E-health Networking, Application Services (HealthCom), Oct 2015, pp. 397–403.

[4] L. S. Piva, A. B. Ferreira, R. B. Braga, and R. Andrade, “falert: An android system for monitoring falls in people with special care,” in Workshop on Tools and Applications of the 20th Brazilian Symposium on Multimedia and Web Systems, 2014.

[5] R. L. Almeida, A. A. Macedo, Í. L. de Araújo, P. A. Aguilar, and R. M. Andrade, “Watchalert: An evolution of the falert app for detecting falls on smartwatches,” in Extended Proceedings of the XXII Brazilian Symposium on Multimedia and Web Systems. SBC, 2016, pp. 124–127.

[6] R. Cucchiara, A. Prati, and R. Vezzani, “A multi-camera vision system for fall detection and alarm generation,” Expert Systems, vol. 24, no. 5, 2007, pp. 334–345.

[7] I. L. de Araújo, L. Dourado, L. Fernandes, R. M. d. C. Andrade, and P. A. C. Aguilar, “An algorithm for fall detection using data from smartwatch,” in 2018 13th Annual Conference on System of Systems Engineering (SoSE), June 2018, pp. 124–131.

[8] A. Ramachandran, A. R., P. Pahwa, and A. K. R., “Machine learning-based techniques for fall detection in geriatric healthcare systems,” in 2018 9th International Conference on Information Technology in Medicine and Education (ITME), Oct 2018, pp. 232–237.

[9] D. Ajerla, S. Mahfuz, F. Zulkernine, and R. C. Pryss, "A real-time patient monitoring framework for fall detection," Wireless Communications and Mobile Computing, vol. 2019, jan 2019, p. 13.

[10] F. Hussain, F. Hussain, M. Ehatisham-ul Haq, and M. A. Azam, "Activity-aware fall detection and recognition based on wearable sensors," IEEE Sensors Journal, vol. 19, no. 12, 2019, pp. 4528–4536.

[11] R. Jansi and R. Amutha, "Detection of fall for the elderly in an indoor environment using a tri-axial accelerometer and kinect depth data," Multidimensional Systems and Signal Processing, 2020, pp. 1–19.

[12] N. Noury et al., "Fall detection-principles and methods," in 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2007, pp. 1663–1666.

[13] C. J. Peñafort-Asturiano, N. Santiago, J. P. Núñez-Martínez, H. Ponce, and L. Martínez-Villaseñor, "Challenges in data acquisition systems: lessons learned from fall detection to nanosensors," in 2018 Nanotechnology for Instrumentation and Measurement (NANOfIM). IEEE, 2018, pp. 1–8.

[14] S.-H. Liu and W.-C. Cheng, "Fall detection with the support vector machine during scripted and continuous unscripted activities," Sensors, vol. 12, no. 9, 2012, pp. 12 301–12 316.

[15] B. Andò, S. Baglio, C. O. Lombardo, and V. Marletta, "A multisensor data-fusion approach for adl and fall classification," IEEE Transactions on Instrumentation and Measurement, vol. 65, no. 9, 2016, pp. 1960–1967.

[16] D. Dziak, B. Jachimczyk, and W. Kulesza, "Iot-based information system for healthcare application: design methodology approach," Applied Sciences, vol. 7, no. 6, 2017, p. 596.

[17] S. Khojasteh, J. Villar, C. Chira, V. González, and E. De La Cal, "Improving fall detection using an on-wrist wearable accelerometer," Sensors, vol. 18, no. 5, 2018, p. 1350.

[18] N. Pannurat, S. Thiemjarus, and E. Nantajeewarawat, "Automatic fall monitoring: a review," Sensors, vol. 14, no. 7, 2014, pp. 12 900–12 936.

[19] A. Ghasemi and S. Zahediasl, "Normality tests for statistical analysis: a guide for non-statisticians," International journal of endocrinology and metabolism, vol. 10, no. 2, 2012, p. 486.

[20] C. Wissler, "The spearman correlation formula," Science, vol. 22, no. 558, 1905, pp. 309–311.

[21] OriginLab, "Origin: Data analysis and graphing software," https://www.originlab.com/Origin, October 2020, (Accessed on 10/23/2020).

[22] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," IEEE Intelligent Systems and their applications, vol. 13, no. 4, 1998, pp. 18–28.

[23] L. Breiman, "Random forests," Machine learning, vol. 45, no. 1, 2001, pp. 5–32.

[24] M. Fernández-Delgado, E. Cernadas, S. Barro, and D. Amorim, "Do we need hundreds of classifiers to solve real world classification problems?" The Journal of Machine Learning Research, vol. 15, no. 1, 2014, pp. 3133–3181.

[25] N. Thai-Nghe, Z. Gantner, and L. Schmidt-Thieme, "Cost-sensitive learning methods for imbalanced data," in The 2010 International joint conference on neural networks (IJCNN). IEEE, 2010, pp. 1–8.

[26] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," Machine Learning, vol. 46, 2002, pp. 389–422.

[27] N. V. C. et. al., "Synthetic minority over-sampling technique," Journal of Artificial Intelligence Research, vol. 16, 2002, pp. 321–357.

[28] M. Hall et al., "The weka data mining software: an update," ACM SIGKDD explorations newsletter, vol. 11, no. 1, 2009, pp. 10–18.

[29] C.-Y. Hsieh, K.-C. Liu, C.-N. Huang, W.-C. Chu, and C.-T. Chan, "Novel hierarchical fall detection algorithm using a multiphase fall model," Sensors, vol. 17, no. 2, 2017, p. 307.

[30] N. M. Josuttis, SOA in practice: the art of distributed system design. " O'Reilly Media, Inc.", 2007.

Pandemics in Hawai'i: 1918 Influenza and COVID-19

Victoria Kala
Department of Mathematics
University of California, Los Angeles
Los Angeles, California
e-mail: vtkala@math.ucla.edu

Katherine Guo
Monte Vista High School
Danville, California
e-mail: katieguo6@gmail.com

Elizabeth Swantek & Alan Tong
Dept. of Mathematics
University of Hawai'i at Manoa
Honolulu, Hawai'i
e-mail: {eswantek,tongalan}@hawaii.edu

Monique Chyba & Yuriy Mileyko & C. Gray
Dept. of Mathematics
University of Hawai'i at Manoa
Honolulu, Hawai'i
e-mail: {chyba,ymileyko,cgray8}@hawaii.edu

Thomas Lee
Office of Public Health Studies
University of Hawai'i at Manoa
Honolulu, Hawaii
https://orcid.org/0000-0002-4588-2348

Alice E. Koniges
Hawai'i Data Science Institute
University of Hawai'i at Manoa
Honolulu, Hawai'i
e-mail: koniges@hawaii.edu

***Abstract*—The COVID-19 pandemic has impacted the health, finances, and lives of citizens across the world. COVID-19 continues to evolve and create "waves" impacting countries at different time points. In this paper, we examine the transmission of COVID-19 between the continental United States and the State of Hawai'i and how that mirrors past influenza pandemics. Firstly, we introduce a summary of the 1918-1920 Influenza Pandemic. We then give an overview of facts and a timeline of government response to COVID-19. In addition, comparisons to the 1918-1920 Influenza Pandemic and 2018-2019, 2019-2020 Influenza seasons are provided. We conclude by addressing open challenges and possible future research directions for studying COVID-19 in Hawai'i.**

***Keywords*—** COVID-19; Influenza; Pandemic 1918; SIR Model.

I. Introduction

The COVID-19 pandemic is far from the first infectious disease that Hawai'i had to deal with. During the 1918-1920 Influenza Pandemic, the Hawaiian islands were not spared as the disease ravaged through the whole world. It is reasonable to ask whether any lessons can be learned from Hawai'i's experience during the 1918-1920 influenza pandemic which could potentially be useful in the fight against COVID-19. It is also sensible to look at the recent influenza activity in the state. While the two diseases are different, mitigation against one of the maladies may prove effective against the other. The goal of this paper is to investigate how the dynamics of COVID-19 pandemic in Hawai'i and its relation to the pandemic in mainland US compares to the previous influenza activity.

In Section II, we present a summary of the 1918-1920 Influenza Pandemic in Hawai'i. In Section III, we present a timeline of government response to and facts of COVID-19 in Hawai'i. In Section IV-B, we discuss COVID-19 and its relation to the 1918-1920 Influenza Pandemic and 2018-2019, 2019-2020 Influenza seasons. The data we present for COVID-19 was gathered from the Centers for Disease Control and Prevention (CDC) [1] and Hawai'i Department of Health (DOH) [2]. The data we present for 2018-2020 Influenza seasons was gathered from the Hawai'i Department of Health [3]. We conclude by addressing challenges and future research for studying COVID-19 in Hawai'i.

II. 1918−1920 Influenza Pandemic

In this section we summarize the findings in [4] discussing the 1918−1920 Influenza Pandemic that killed more than 21 million people, including over 675,000 Americans and more than 2,300 people in Hawai'i. The global timeline of this worldwide pandemic can be seen in Fig. 1. There are three types of influenza (A, B, C), of which Influenza A is considered the type of virus that produced the 1918 epidemic in Hawai'i. The first cases seen on Oahu were at the island's military and naval bases at the end of June 1918. The first wave of the 1918 pandemic quickly rolled over Oahu, lasting through July and into August. The second wave occurred December and January. There were severe shortages of physicians due to World War I, along with shortages of hospital beds, nurses, and other medical personnel and facilities. Although mainland communities had closed schools, theaters, and churches as public health measures during the pandemic, Island officials resisted similar measures arguing the closure would have little or no effect on death rates. Eventually some businesses and institutions temporarily closed, and indoor public gatherings were prohibited.

Statistics of the pandemic are difficult to find. Morbidity statistics were largely unavailable before October 21, 1918, on which date influenza became a reportable disease. However, the authors do have morbidity statistics before this date as seen in Fig. 2, Fig. 3, and Table I. Influenza morbidity has typically been subject to proportionately more under reporting than flu mortality. Many states were not yet part of the U.S. death registration area. Hawai'i, a territory at the time, was omitted from national totals and reported in separate tables until being admitted to the death registration area in 1917. Another statistical problem is that morbidity figures generally omit cases of pneumonia, which were often the ultimate outcome for influenza patients.

Fig. 2 shows the death rates for the U.S. death registration area and the territory of Hawai'i for 1917-1920. Note that Hawai'i's flu-related deaths peaked in 1920, two years after the mainland; the reasons for this lag remain unclear.

Death rates were highest for children under 5 years of age as seen in Table. I, lowest for those between five and nineteen. Flu deaths were exceptionally high for Japanese and pure Hawaiians as seen in Fig. 3 and Table I while male and female rates were about the same.

III. COVID-19

We summarize some important facts and events of COVID-19 in Hawai'i. A detailed timeline can be found in Fig. 4. The first confirmed COVID-19 case was reported on March 6, 2020 [2]. As of September 22, 2020, there have been 11,522 COVID-19 cases, with the majority of them occurring during a large second wave, starting in the first week of June and peaking mid August. During the first

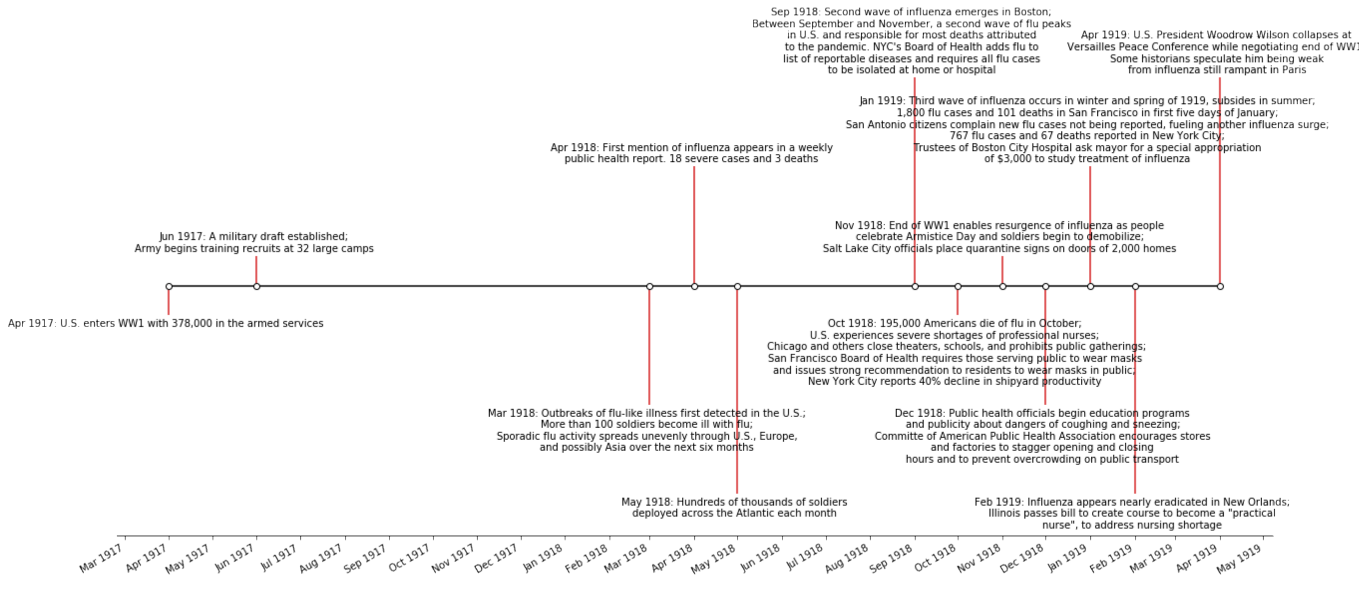

Fig. 1. Influenza (flu) Pandemic Events and Dates 1917-1919

wave of COVID-19 cases, the major contributing factor was due to tourists, prompting the shut down of air and sea travel. However, the overwhelming cause of cases during the second wave can be attributed to wide community spread [5]. As of September 22, 2020, there has been 749 hospitalizations and 120 deaths from the disease [2], a death rate of approximately 1.04%. Fig. 5 shows a plot of average daily cases and deaths.

We have listed the COVID-19 deaths by age, sex, and race in Table II. COVID-19 death victims tend to be older in age, and it also appears to affect males more than females. The DOH data for race breakdown seems to be incomplete, as the total number of deaths categorized by race do not add up to the total number of deaths reported at that time (total of 119). We have included Native Hawaiian, Chinese, Black, and Other in the “Unreported” category as groups with fewer than 5 deaths are not reported. We have also plotted the race breakdown as seen in Fig. 6. With the information given, we can observe that Pacific Islanders are disproportionately affected by COVID-19 even though they make up only about 4% of Hawai’i population. On the counterpart, data suggests that Native Hawaiians are less impacted.

Fig. 7 displays the daily new cases as well as a fit obtained using a standard discrete compartmental model. This model divides the population into four compartments: Susceptible (not currently infected), Exposed (infected with no symptoms), Infected (infected with symptoms), Removed (recovered or deceased), i.e., an SEIR model

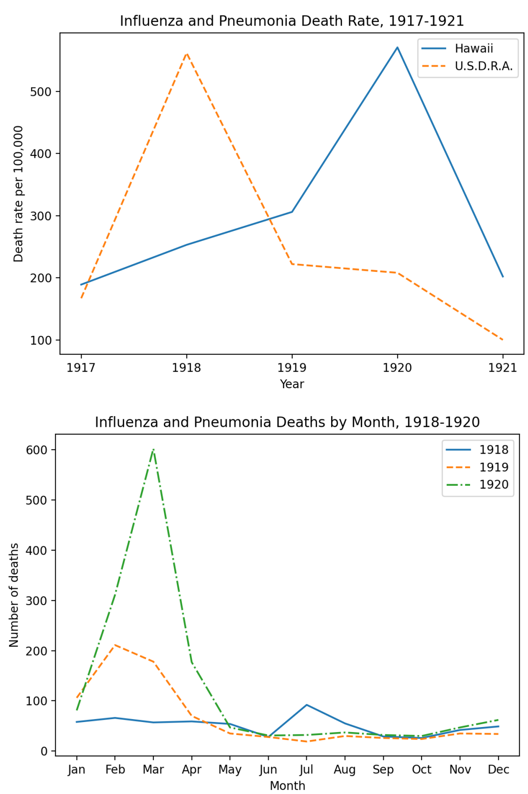

Fig. 2. Up: Deaths from Influenza and Pneumonia for Hawai’i, 1918-1920. Bottom: Comparing Death for Hawai’i and U.S. death registration area. Data available from [4].

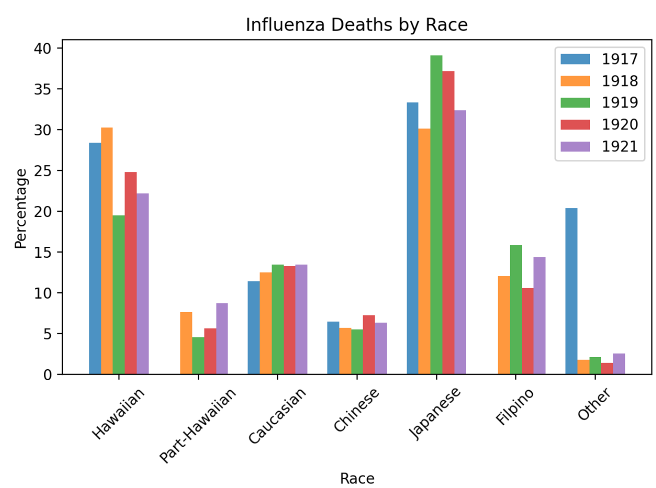

Fig. 3. Deaths from Influenza and Pneumonia by Race for Hawai’i, 1917-1921. Part-Hawaiians and Filipinos were included in “Other” for 1917. Data available from [4].

Copyright (c) IARIA, 2020. ISBN: 978-1-61208-817-4

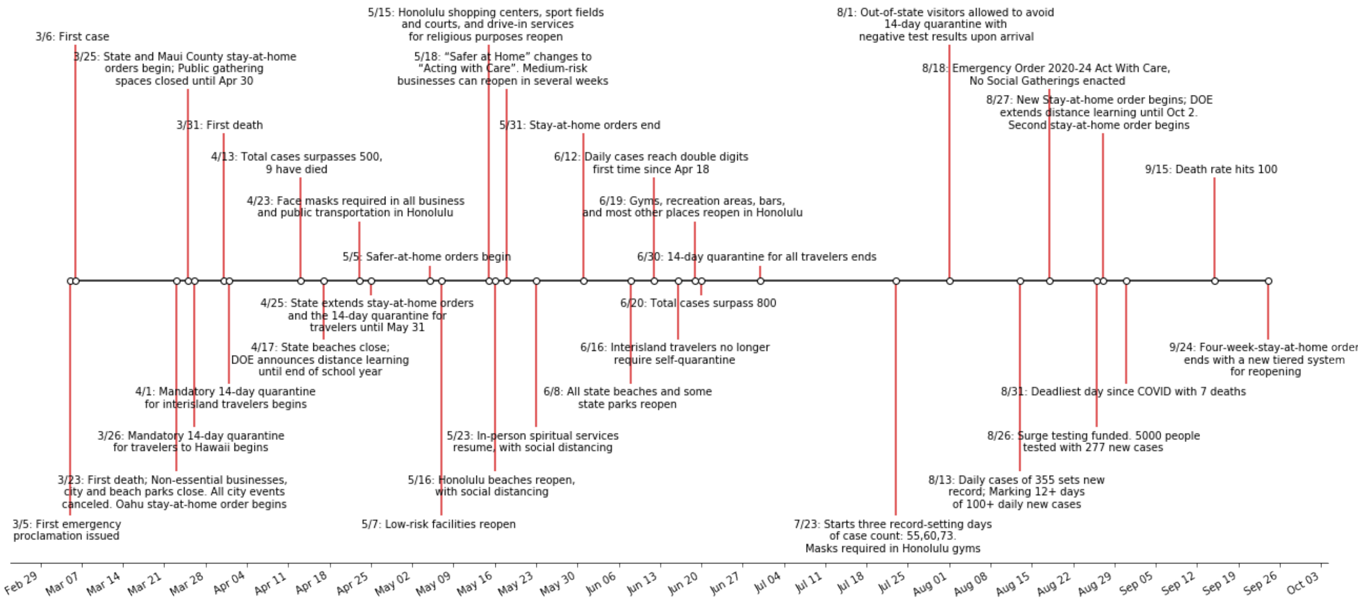

Fig. 4. COVID-19 Hawai'i Events and Dates 2020

(see e.g., [6], [7]). Denoting the populations in these compartments at time t by $S(t)$, $E(t)$, $I(t)$, $R(t)$, respectively, the dynamics can be described by the following equations:

$$S(t+1) = e^{-\lambda(t)} S(t) \tag{1}$$
$$E(t+1) = (1 - e^{-\lambda(t)}) S(t) + (1-p)E(t) \tag{2}$$
$$I(t+1) = pE(t) + (1-r)I(t) \tag{3}$$
$$R(t+1) = rI(t) + R(t), \tag{4}$$

where $\lambda(t)$ is the hazard rate that is computed as $\lambda(t) = \beta(I(t) + \varepsilon E(t))$. Here, β is the baseline transmission rate, ε is the transmission reduction factor for asymptomatic transmission, p is the rate for onset of symptoms and r is the removal (recovery or death) rate. The transmission rates for the fit are described in Table III. Fig. 5 displays on the same graph both the daily cases and the death to highlight the delay of the onset of deaths compare to the increase in daily cases.

TABLE I
INFLUENZA AND PNEUMONIA DEATHS BY AGE, SEX, AND RACE IN HAWAI'I, 1917-1921[a]

Subject	1917	1918	1919	1920	1921
Total	447	615	796	1,489	550
Age					
Under 5 years	294	360	274	482	364
5 to 19 years	24	34	86	146	27
20 to 39 years	38	96	86	146	27
40 to 59 years	44	74	112	247	57
60 years and over	47	50	54	85	36
Age	–	–	–	2	1
Sex[b]					
Male	273	346	440	940	–
Female	174	269	356	649	–
Race[c]					
Hawaiian	127	187	155	369	122
Part-Hawaiian	–	47	36	84	48
Caucasian	51	77	107	197	74
Chinese	149	185	311	553	178
Filipino	–	74	126	157	79
Others	91	11	17	21	14

[a] Data available from [4].
[b] Sex was not recorded for 1921.
[c] Part-Hawaiians and Filipinos combined with "Other" in 1917.

Next, we compare the situation in Hawai'i to the entire United States. As of September 22, 2020 the United States has had a total of 199,462 deaths out of 6,825,697 cases [1]. The death rate in Hawai'i is approximately 1.04%, which is lower than the current national average of approximately 2.92%. Fig. 8 represents the daily cases and death per capita for the State of Hawai'i compared to the United States. Note that the data collected for the US also includes Hawai'i data. A delay can be observed between the US and Hawai'i peaks in the number of cases. The number of deaths appears to be rising for Hawai'i.

IV. DISCUSSION

A. *1918-1920 Influenza Pandemic versus COVID-19*

It may be too early to compare the 1918-1920 Influenza Pandemic and COVID-19 Pandemic, and data for the 1918-1920 Influenza Pandemic are limited to deaths and not daily cases, whereas in the

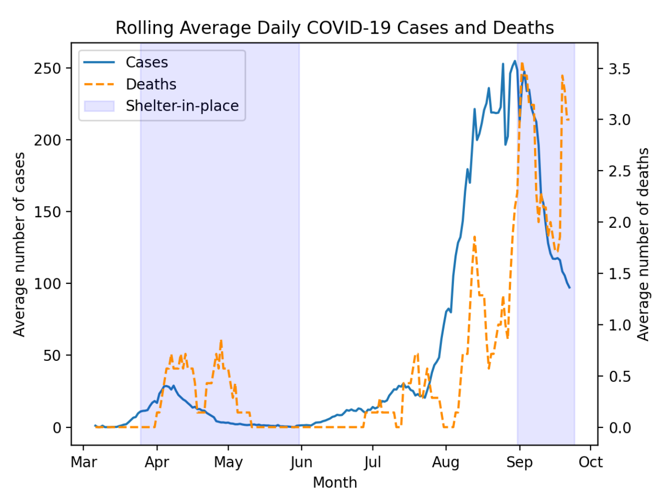

Fig. 5. 7-day rolling average daily cases and deaths for Hawai'i. Data available from the CDC.

Copyright (c) IARIA, 2020. ISBN: 978-1-61208-817-4

TABLE II
COVID-19 DEATHS BY AGE, SEX, AND RACE[a]

Age	Deaths	Sex	Deaths	Race	Deaths	State Population
30-39 years	1	Male	79	Caucasian	8	25%
40-49 years	5	Female	40	Native Hawaiian	<5	21%
50-59 years	12	Total	119	Pacific Islander	21	4%
60-69 years	19			Filipino	19	16%
70-79 years	37			Japanese	18	15%
80+ years	45			Chinese	<5	4%
Total	119			Other Asian	7	4%
				Black	<5	2%
				Other	<5	8%
				Unreported	46	
				Total	119	

[a] Data available from the Hawai'i DOH up until September 18, 2020.

case of COVID-19 pandemic we have access to much more complete data. We do however note some similarities and differences between the two pandemics.

First, let us compare the deaths using Tables I and II and Fig.s 3 and 6. Current data show that the demographic perspectives on the mortality of 1918-1920 Influenza Pandemic and COVID-19 Pandemic differ with the early pandemic affecting younger individuals under the age of 5 while COVID-19 is almost exclusively affecting individual older than 60 years old. Both pandemic seem to be dominantly impacting male versus female. In terms of race, there is some variations due to a different racial distribution of the individual making up the Hawai'i population, however in both cases socio-economic minorities are being disproportionately affected. Hawaiian and Japanese (then a minority) made a large percentage of deaths in the 1918-1920 influenza pandemic, while COVID-19 is dominantly affecting Pacific Islanders. Japanese also make up a large percentage of deaths for the COVID-19 pandemic, although it appears that other groups are at higher proportion versus Japanese groups, e.g., Pacific Islanders.

An interesting comparison between 1918-1920 Influenza Pandemic and COVID-19 Pandemic is the delay of the waves between the US and Hawai'i. It can be clearly seen from Fig. 2 and Fig. 6. The delay in the 1918 pandemic is much larger and we suspect it is due to the fact that traveling was more limited and tedious but the delay still exists today (few months). It is too early to know how the curves will evolve for COVID-19 and something that we will be following.

B. 2018-2020 Influenza Seasons versus COVID-19

At the beginning of the pandemic, COVID-19 was often dismissed as another variant of influenza. For the second part of our discussion, we attempt to compare and contrast data between COVID-19 and influenza, since both data tables are rapidly available to the public. We utilize the 2018-2019 influenza season as the control, while comparing it to both COVID-19 and and the 2019-2020 influenza season.

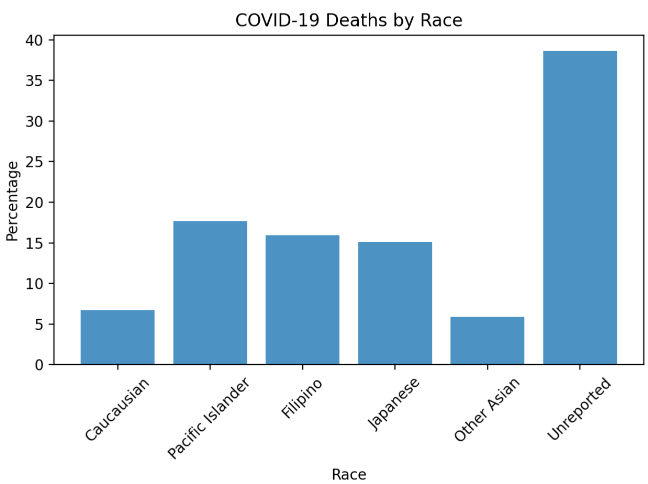

Fig. 6. COVID-19 deaths by race. Native Hawaiian, Chinese, Black, and Other races are included in the "Unreported" category. Data available from the Hawai'i DOH.

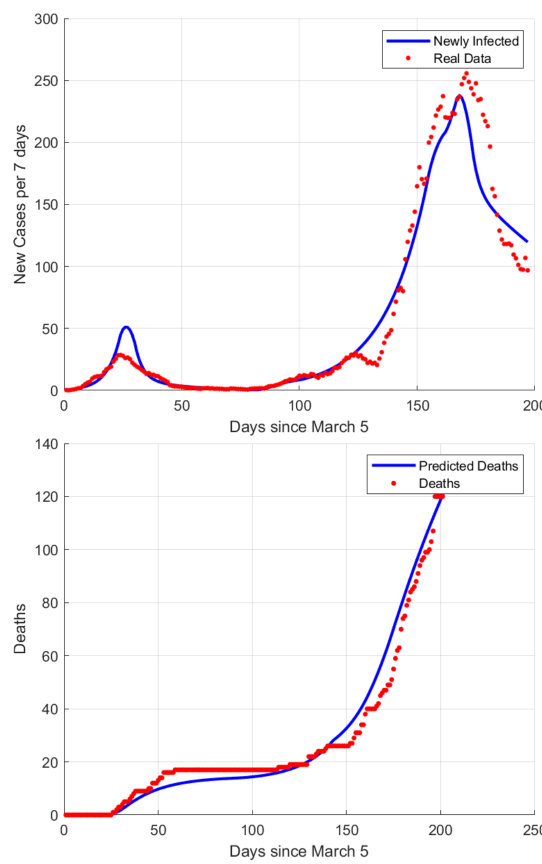

Fig. 7. Up: 7 days rolling daily cases. Dots are the actual data and the plain line represents the model. Bottom: Cumulative death cases as well as the fit from the model.

TABLE III
TRANSMISSION RATES FOR THE MODEL

Changes in the Value of β.						
Mar 6	Apr 2	May 20	May 30	Jun 10	Aug 11	Aug 23
β=0.43	β=0.02	β=0.01	β=0.21	β=0.15	β=0.11	β=0.05

Copyright (c) IARIA, 2020. ISBN: 978-1-61208-817-4

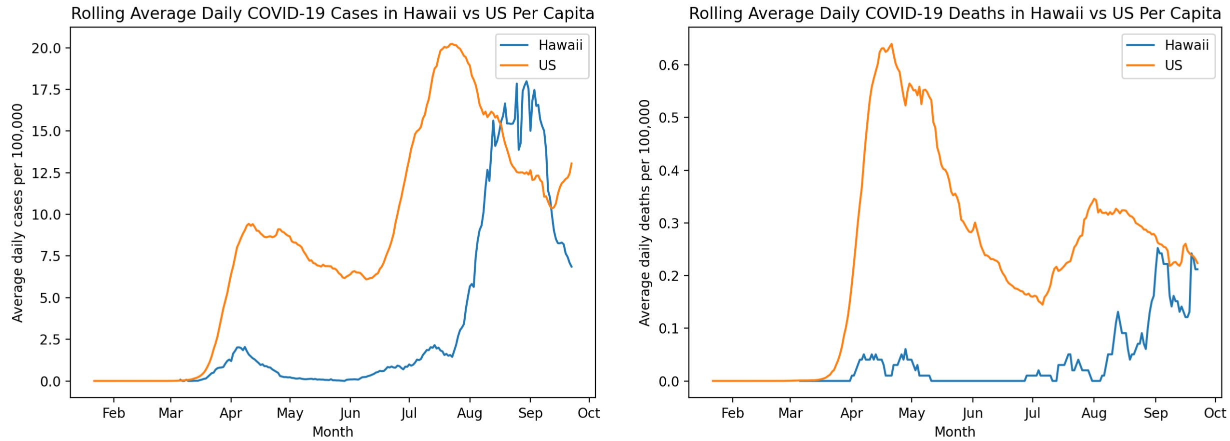

Fig. 8. The 7-day rolling average for daily cases (left) and deaths (right) in Hawai'i vs. the US. Date available from the CDC.

COVID-19 is mostly transmitted via airborne droplets, similarly to Influenza [8]. However, additional research has shown COVID-19 can be transmitted via aerosol [9].

Historically, a major contributor to influenza season in Hawai'i is due to tourists escaping winter, explaining why Hawai'i does not have a traditional flu season compared to the rest of the United States.

As seen in Fig. 9, the 2019-2020 influenza season was very similar to the 2018-2019 influenza season up until the shelter-in-place order for COVID-19, in which the number of cases for the 2019-2020 influenza season plummets significantly. COVID-19 cases similarly decreased during the shelter-in-place order. Further, the number of weekly cases for the 2019-2020 influenza appear to not rise after the shelter-in-place order was lifted, unlike the COVID-19 cases which saw a drastic increase. The peak in number of weekly cases for COVID-19 is more than double of the peaks for the influenza seasons. The reasons for this observations are unclear, it can be due to confirmation bias due to surge testing and other aggressive testing measures for COVID-19 while influenza testing is limited to surveillance and individual seeking medical help for influenza like illness. It can also be due to longer infectious period for COVID-19 compare to influenza. Note that as it was observed when comparing the 1918 pandemic to COVID-19, the two diseases target different age demographic: the influenza virus targets individuals under 20, while COVID-19 targets individuals over 20.

We now direct our attention to testing for the influenza seasons and COVID-19. The 2018-2019 influenza had an average positivty rate of 17%, the 2019-2020 influenza has an average positivity rate of 13.7%, and COVID-19 has an average positivity rate of 2.2%. Fig. 10 reflects the effect of the COVID-19 pandemic as it forces stronger testing. The increase in testing resulted in more negative tests for the 2019-2020 influenza season versus the 2018-2019 influenza season. In Fig. 11, the number of tests for the 2019-2020 influenza season increased significantly before the shelter-in-place order, but then decreased during shelter-in-place. We are not sure of the reason for this but one hypothesis could be that those exhibiting influenza like illness avoided their PCP out of fear from contracting COVID-19 and therefore did not get tested. Finally, Fig. 12 displays the counts for weekly death for the 2018-2019 and the 2019-2020 influenza as well as COVID-19 for the state of Hawai'i. The shelter in place did occur at a time the number of death for influenza in 2018-2019 was decreasing anyway and we see a similar behavior for influenza 2019-2020 with slightly less cases but striking is the different behavior for the weekly count of deaths for COVID-19. It is extremely hard at this stage to anticipate how influenza and COVID-19 are going to interact in the coming Fall and Winter.

V. Conclusion and Future Work

This paper depicts the current situation in the State of Hawai'i regarding the current COVID-19 pandemic. This data analysis needs to be pursued further to continue monitoring the evolution of the spread of COVID-19. It is still very early with the new pandemic to draw definitive conclusions, but there are some similarities that seem

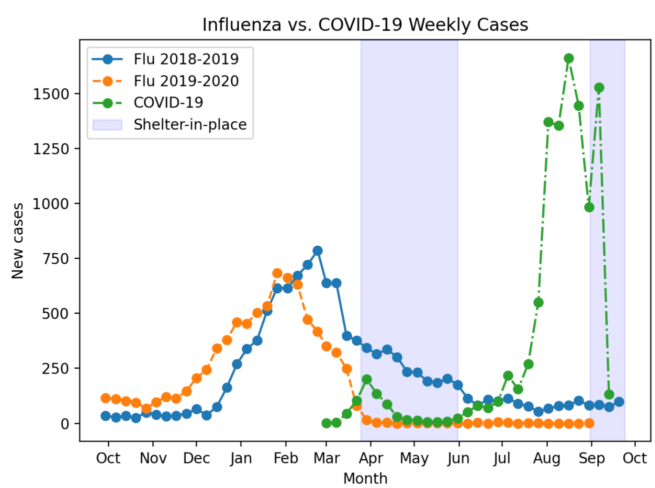

Fig. 9. Weekly cases for influenza and COVID-19. Each dot represents the weekly case. Data available from the Hawai'i DOH.

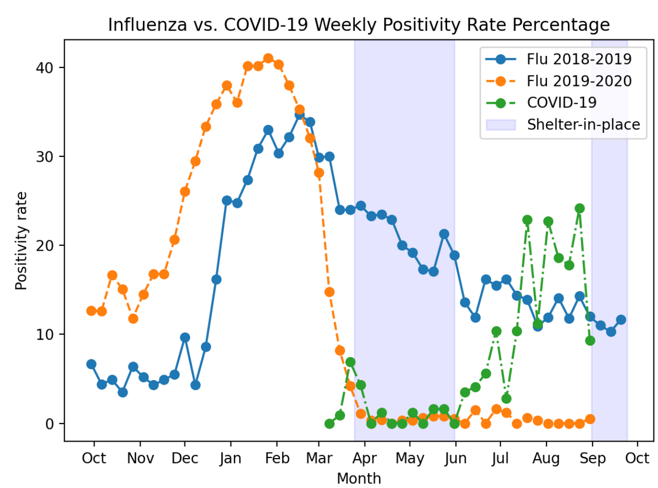

Fig. 10. Positivity rate percentages for influenza and COVID-19. Each dot represents the weekly rate. Data available from the Hawai'i DOH.

Copyright (c) IARIA, 2020. ISBN: 978-1-61208-817-4

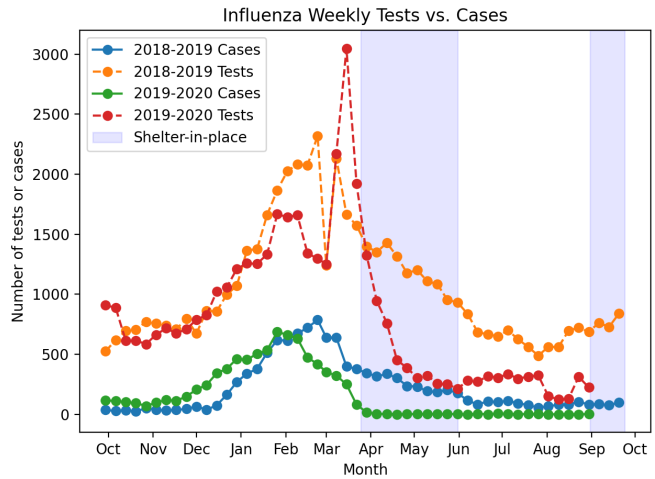

Fig. 11. Weekly influenza cases along with reported testing numbers. Each dot represents the weekly number of tests or cases. Data available from the Hawai'i DOH.

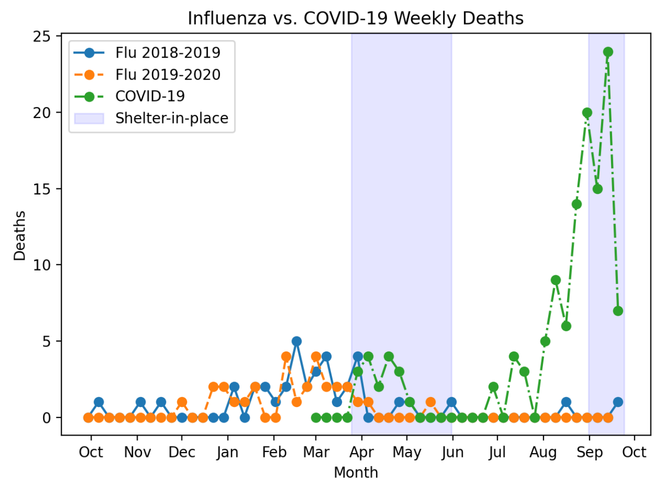

Fig. 12. Weekly deaths from influenza and COVID-19. Each dot represents the weekly number of deaths. Data available from the CDC.

to be highlighted with the data. Both cases (1918 and today) rely on non-pharmaceutical measures to contain the pandemic, at least for COVID-19 until a vaccine will be available. Transportation being much improved, especially over long distances, COVID-19 is difficult to contain. In 1918, it was World-War that was a major trigger for the spread of the disease, while now it is our lifestyle. In both cases, we see Hawai'i lagging behind the US mainland with regards to the waves of deaths related to the diseases.

There is much work to be done to understand the spread of the disease in Hawai'i. Moreover, with the State ready to reopen for travelers, and to have students going back to school in person, the next few weeks will be critical for the evolution of COVID-19. Influenza pandemic of 1918 exhibited three main waves spread over a three years period, and the primary concern for COVID-19 right now is an upcoming second wave.

From this investigation, we conclude that COVID-19 acts differently qualitatively and quantitatively from influenza virus; the danger of COVID-19 should be viewed seriously and not dismissed.

ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation under Grant No. 2030789. We would like to thank Michael Jones, retired faculty from the Physics department at the University of Hawai'i Manoa for his help in this work and data sharing.

REFERENCES

[1] Center for Disease Control and Prevention, *CDC COVID Data Tracker*, Center for Disease Control and Prevention. [Online]. Available from: https://covid.cdc.gov/covid-data-tracker/. [Retrieved: Sep. 2020].

[2] State of Hawai'i Department of Health Disease Outbreak Control Division, *COVID-19: Current Situation in Hawai'i*, State of Hawai'i Department of Health Disease Outbreak Control Division. Available from: https://health.hawaii.gov/coronavirusdisease2019/what-you-should-know/current-situation-in-hawaii/ [Retrieved: Sep. 2020].

[3] State of Hawai'i Department of Health Disease Outbreak Control Division, *Influenza (Flu)*, State of Hawai'i Department of Health Disease Outbreak Control Division. [Online]. Available from: https://health.hawaii.gov/docd/disease_listing/influenza-flu/. [Retrieved: Sep. 2020].

[4] R. C. Schmitt and E. C. Nordyke. "Influenza Deaths in Hawai'i, 1918-1920." 1999.

[5] A.M. Ollstein and D. Goldberg, *Paradise lost: How Hawai'i went from Covid-19 star to cautionary tale*, Politico, Aug. 2020. [Online]. Available from: https://www.politico.com/news/2020/08/30/hawaii-coronavirus-cautionary-tale-404564. [Retrieved: Sep. 2020].

[6] W. M. Getz and E. R. Dougherty. "Discrete stochastic analogs of Erlang epidemic models," *Journal of biological dynamics*, vol. 12, no. 1, pp. 16-18, 2018.

[7] M. J. Keeling and P. Rohani, *Modeling infections diseases in humans and animals*, Princeton University Press, 2011.

[8] Center for Disease Control and Prevention, *COVID-19 FAQ*, Center for Disease Control and Prevention. [Online]. Available from: https://www.cdc.gov/coronavirus/2019-ncov/faq.html. [Retrieved: Sep. 2020].

[9] J.H. Tanne, *Covid-19: CDC publishes then withdraws information on aerosol transmission*, BMJ 2020;370:m3739. [Online]. Available from: https://www.bmj.com/content/370/bmj.m3739. [Retrieved: Sep. 2020].

Epidemiological model of the spread of COVID-19 in Hawaii's challenging fight against the disease

Monique Chyba, Yuriy Mileyko, Oleksandr Markovichenko, Richard Carney
Dept. of Mathematics
University of Hawai'i at Manoa
Honolulu, Hawai'i, United States
{chyba,ymileyko,markov,rscarney}@hawaii.edu

Alice E. Koniges
Hawai'i Data Science Institute
University of Hawai'i at Manoa
Honolulu, Hawai'i, United States
koniges@hawaii.edu

Abstract*—Hawai'i and similar island populations can follow a different course of pandemic spread than large cities/states/nations and are often neglected in major studies. We provide a detailed epidemiological model of the spread of COVID-19 in Hawai'i and explore effects of different intervention strategies in both a prospective and retrospective fashion. We use a modified compartmentalized extended Susceptible-Exposed-Infected-Recovered (SEIR) model with a simple step function time dependence calibrated using the current data. We model asymptomatic carriers and actual mitigation strategies such as social distancing, contact tracing and quarantine policy. We find different outcomes for different scenarios: predicting retrospectively that if we successfully isolated 52% (alternately 38.8%) of the asymptomatic on days 2-4 (alternately 3-5) after exposure as we came out of the first lockdown, we would have reduced daily cases, hospitalisation and Intensive Care Unit (ICU) occupancy by about 74% (alternately 50%), where alternately refers to the second of two retrospective modeling scenarios. Going forward, we forecast that successful identification and isolation of 49% between days 2-4 of exposure as well as compliance from individuals to reduce the transmission rate by 14.4% provides a scenario where the daily cases would peak for a third time at a moderate triple digit rate of 193 in early December, potentially avoiding a third lockdown. Furthermore, because of the unique isolation of Hawai'i to incoming travelers fitting our specific data and using our methodology to predict outcomes serves as an important semi-controlled experiment to help others in applying epidemiological models to their populations.

***Keywords*—** COVID-19; Compartmentalized Epidemiological Model; Contact Tracing; Pandemic Mitigation; Hawai'i.

I. INTRODUCTION

Hawai'i and other US Islands have recently been noted by the media as COVID-19 hotspots after a relatively calm period of low case rates. U.S. Surgeon General Jerome Adams came in person on August 25 to Oahu to address the alarming situation. In this paper, we capture the peculiarity of the situation in Hawai'i and provide detailed modeling of current virus spread patterns aligned with dates of lockdown and similar measures. We use this analysis to formulate predictive scenarios.

Hawai'i finds itself in a unique position due to its extremely isolated geographic location, mostly linear population distribution along the coast, and a heavy dependence on the tourism and hospitality sectors of the economy. While the first two factors appeared advantageous in the fight against the disease, the latter one creates a tempering effect on feasible long-term mitigation efforts, since too stringent an approach may lead to a catastrophic impact on the economy. We study the unique aspects of Hawai'i from both a social and data-driven modeling perspective to understand and recommend the critical intervention measures that make the most impact on spread of the disease while mitigating societal adversities.

This is not the first Hawai'i encounter with an invasive virus. In the past, measles, whooping cough, dysentery, and influenza decimated the native Hawaiian population. In the first recorded introduction of major diseases to the islands, measures were taken to prevent sailors from being in close contact with natives; ultimately, this failed due to complications demanding sailors stay on the islands for several days. Upon later visits, the westerners noticed the heavy impact of the disease upon the entire archipelago. Later on in the Kingdom's history, foreign ships brought about more diseases: cholera (1804), influenza (1820s), mumps (1839), measles and whooping cough (1848-9) and smallpox (1853). King Kamehameha V, noticing the loss of his people, established the Kingdom's first quarantine station, which later moved from Honolulu to Moloka'i. Not only were a large portion of Native Hawaiian lives lost, but schools were left empty and the economy disrupted due to the great mortality [1]–[3]. Today, Hawaii remains vulnerable to disease because it is geographically isolated and does not have hospital and other facilities capable of treating large numbers of infected non-residents, nor does it have the ability to shift excess COVID-19 cases to neighboring hospitals or care centers other than perhaps military ships.

Today facing the COVID-19 pandemic, the state government has taken several measures to mitigate spread, including a stay-at-home order and an incoming arrival 14-day quarantine, a reflection of the past. As with minority populations in other states, certain populations (in our case self-identified Pacific Islanders exclusive of native Hawaiians) are impacted at a significantly larger rate, both in terms of testing positive for the disease compared to their population percentage and presumably with respect to the economical loss since this population depends heavily on the tourist industry for economic stability [4] [5]. The March stay-at-home order brought applause when the epidemic was stomped flat but as a result Hawai'i remained extremely vulnerable to the disease exemplified by the current alarming situation in which the islands saw a very significant second wave of infections. The state's seven day average case rate per 100,000 of populations went from months at the bottom of the US list to holding a clear spot in the top 15 as of the ending days of August 2020 [6].

Compartmentalized SEIR models of the COVID-19 provide the basis for much of the current epidemiological modeling efforts worldwide, however variants in the compartmental choices and corresponding variables allow for parameter matching and optimizations, thus providing useful predictive information specific to our Island population [7] [8]. In this paper, we adapt and use these models on Oahu data. Oahu is the most populated island in the chain, providing an appropriate data set for interpretation of our models as well as guidance for the entire state. We focus specifically on Oahu, the most-affected by COVID-19 Island as of now, since each island (or group of islands in the case of Maui) has its own mayor and thus restrictions

Copyright (c) IARIA, 2020. ISBN: 978-1-61208-817-4

and governmental actions may vary slightly within the entire state as they are determined not only uniformly by the Governor but also by the Mayors and local governments of the outer islands.

Our simulations demonstrate that to control the spread of COVID-19 both actions by the State in terms of testing, contact tracing and quarantine facilities as well as individual actions by the population in terms of behavioral compliance to wearing a mask and gathering in groups are vital. They also explain the turn for the worst Oahu took after a very successful stay-at-home order back in March.

It is time critical that models for COVID-19 transmission are applied across a variety of US regions with populations exhibiting different characteristics. Results are presented for Italy [9] and Austin, Texas [10] highlighting the impact of proper timing for efficient mitigation measures. Tuning our model to Hawai'i specific data ensures proper assessment of the local spread of the disease that may not be feasible via larger and more common studies. Also, the controlled environment plays a role in assessing the effectiveness of current and future mitigation strategies. Our methodology is directly applicable to other States and counties with similar accessible data. Observed qualitative behavior of the transmission rate for the States of Alaska, Oregon and Montana, as well as the corresponding numbers for the daily cases, are comparable to Hawai'i's [11] [12]. Our work suggests that a lack of individual compliance prevented a significant decrease in the number of daily cases in these states, and indicates that they are likely to see the daily case numbers increase further following the recent increase in transmission rate. There is a need to either expand asymptomatic isolation or request more severe compliance from the population.

II. MATHEMATICAL MODEL AND PARAMETERS

A. SEIR Compartmentalized Model

To model the spread of COVID-19, we employ a compartmentalized model inspired by [13], which is based on a standard discrete SEIR model. As in the standard SEIR model, we partition a given population into four compartments: Susceptible (not currently infected), Exposed (infected with no symptoms), Infected (infected with symptoms), Removed (recovered or deceased). However, to better capture the dynamics of the infection, we divide the whole population into two population groups: the general community and healthcare workers (healthcare workers play a vital role and are exposed in different ways than the general community, see for instance [14] [15]). We shall denote these groups by C and H, respectively. These groups interact with each other, and each of them consists of the aforementioned compartments. Hence, variables representing the compartments may be decorated with a sub-index c or h, to indicate the appropriate group. In addition, compartments Exposed and Infected (in each population group) are split into multiple stages to better reflect the progression of the disease.

The dynamics of each population group have two distinguished parts: the dynamics of Susceptible individuals, and the dynamics of the rest of the compartments. The former is governed by the *hazard rate*, $\lambda(t)$, which depends on time and essentially determines the probability, $1 - e^{-\lambda(t)}$, of an individual becoming exposed at time t. The hazard rate is different for different population groups and takes into account interactions between the groups, thus coupling their dynamics.

The equations for the dynamics of the two population groups are essentially the same and are given below. Only the hazard rate and the parameters determining transition rates into quarantine may be different between the two groups.

$$S(t+1) = e^{-\lambda(t)} S(t) \tag{1}$$

$$E_0(t+1) = (1 - e^{-\lambda(t)}) S(t) \tag{2}$$

$$E_i(t+1) = (1 - p_{i-1})(1 - q_{a,i-1}) E_{i-1}(t), \quad i = 1, \dots, 13 \tag{3}$$

$$E_{q,i}(t+1) = (1 - p_{i-1})(q_{a,i-1} E_{i-1}(t) + E_{q,i-1}(t)), \quad i = 1, \dots, 13 \tag{4}$$

$$I_0(t+1) = \sum_{i=0}^{13} p_i (1 - q_{a,i}) E_i(t) \tag{5}$$

$$I_1(t+1) = (1 - q_{s,0}) I_0(t) \tag{6}$$

$$I_2(t+1) = (1 - q_{s,1}) I_1(t) + (1 - r)(1 - q_{s,2}) I_2(t) \tag{7}$$

$$I_j(t+1) = r(1 - q_{s,j-1}) I_{j-1}(t) + (1 - r)(1 - q_{s,j}) I_j(t), \quad j = 3, 4 \tag{8}$$

$$I_{q,0}(t+1) = \sum_{i=0}^{13} p_i (q_{a,i} E_i(t) + E_{q,i}(t)) \tag{9}$$

$$I_{q,1}(t+1) = I_{q,0}(t) + q_{s,0} I_0(t) \tag{10}$$

$$I_{q,2}(t+1) = I_{q,1}(t) + q_{s,1} I_1(t) + (1 - r)(q_{s,2} I_2(t) + I_{q,2}(t)) \tag{11}$$

$$I_{q,j}(t+1) = r(q_{s,j-1} I_{j-1}(t) + I_{q,j-1}(t)) + (1 - r)(q_{s,j} I_j(t) + I_{q,j}(t)), \quad j = 3, 4 \tag{12}$$

$$R(t+1) = R(t) + r I_4(t) + r I_{q,4}(t) + (1 - p_{13}) E_{13}(t) + (1 - p_{13}) E_{q,13}(t) \tag{13}$$

Below is a detailed description of the variables, all of which depend on time, t, measured in days.

- **Variable** $S(t)$**.** The number of susceptible individuals.
- **Variables** $E_i(t)$**.** The number of asymptomatic infected individuals i days after exposure who are not quarantined.
- **Variables** $E_{q,i}(t)$**.** The number of quarantined asymptomatic infected individuals i days after exposure.
- **Variables** $I_j(t)$**,** $i = 0, 1$**.** The number of symptomatic infected individuals i days after the onset of symptoms who are not quarantined.
- **Variables** $I_j(t)$**,** $j = 3, 4, 5$**.** The number of symptomatic infected individuals at the nominal stage i of the illness. Note that a person can stay at a given stage for several days.
- **Variables** $I_{q,j}(t)$**,** $j = 0, 1$**.** The number of quarantined symptomatic infected individuals, with j representing either the number of days after the onset of the symptoms ($j = 0, 1$), or the stage of the illness ($j = 2, 3, 4$).
- **Variable** $R(t)$**.** The number of removed (recovered or deceased) individuals.

Splitting exposed individuals into multiple stages, E_i, allows us to capture possible differences in the progression of the asymptomatic phase of the disease. Importantly, it allows us to take into account that, according to the Centers for Disease Control and Prevention (CDC) as well as other sources, about 40% of people who contract SARS-CoV-2 remain asymptomatic, and the incubation period for those who do develop symptoms is somewhere between 2 to 14 days after exposure, with the mean incubation period between 4 and 6 days [16]–[18]. Individuals who do not develop symptoms after 14 days are assumed recovered. The use of the quarantine sub-compartments,

Copyright (c) IARIA, 2020. ISBN: 978-1-61208-817-4

$E_{q,i}$, allows us to capture the effect of contact tracing and the reduced transmission rate for quarantined individuals.

Similarly, having multiple stages for infected individuals better reflects progression of the symptomatic phase of the disease. The first two stages represent the first two days of symptoms, but the next three should be understood as phases of the immune system fighting the disease. There is a substantial variability (due to age as well as other factors) in the number of days any given person can spend at each stage. Our model implicitly assumes that the symptomatic phase of the illness lasts at least 5 days (in the unlikely case that each stage lasts just one day). From the variables above, we can compute hospitalization count and the total number of ICU beds (it is assumed 8% of people with symptoms are hospitalized, and of those, 20% enter the ICU.)

As we mentioned, a crucial part of the dynamics relates to the hazard rate. For the general community, group C, we have

$$\lambda_c(t) = \beta\Big[(I_c + \varepsilon E_c) + \gamma((1-\nu)I_{c,q} + \varepsilon E_{c,q}) + \rho[(I_h + \varepsilon E_h) + \gamma((1-\nu)I_{h,q} + \varepsilon E_{h,q})]\Big]/N_c, \quad (14)$$

where we suppressed the dependency on t on the right for convenience. Subscripts c and h indicate the general community and healthcare workers, respectively, and subscript q indicates quarantined individuals. Variables E and I here represent the sum over all the stages within these compartments. N_c denotes the mixing pool for the general community, computed as

$$N_c(t) = S_c + E_c + I_c + R_c + \rho(S_h + E_h + I_h + R_h). \quad (15)$$

For the healthcare worker group, we have

$$\lambda_h(t) = \rho\lambda_c + \beta\eta\Big[(I_h + \varepsilon E_h) + \kappa\nu(I_{h,q} + I_{c,q})\Big]/N_h, \quad (16)$$

where $N_h(t) = S_h + E_h + I_h + R_h$.

B. Parameters and Initial Conditions

The model parameters have been chosen to reflect the spread of COVID-19 on Oahu, taking into account the fit for the daily cases and active hospitalizations. Most parameters remain fixed over time. However, parameter β, capturing the basal transmission rate due to various interactions among individuals, as well as parameters reflecting the rates of quarantine can change over time along with active mitigation measures. Specifically, we use several different values of β that capture changes in COVID-19 policy on Oahu (specific values can be seen in Table II). Table I provides the description of the model parameters and their values.

The values of the basal transmission rate β change over time, as shown in Table II, and are obtained by minimizing the sum of squared differences between the recorded number of daily cases and the number given by the model (the latter is calculated as the number of individuals about to be quarantined). The optimization is done using the Levenberg–Marquardt algorithm [19], which is well suited for solving the nonlinear least squares problem. The Jacobian of the objective function, required by the algorithm, is computed exactly by implementing the dynamics of the derivatives of our variables with respect to β.

It is useful to regard parameters p_i, $q_{a,i}$, and $q_{s,i}$ as probabilities, which is their role in stochastic SEIR epidemiological models [20]. The values of p_i are chosen to reflect the observations that, if symptoms do develop, it takes between 2 to 14 days, with a mean between 4 and 6 days [16]. We also take into account the estimate that about 40% of all infections remain asymptomatic. In the stochastic setup, the probability to remain asymptomatic is given by $\prod_{i=0}^{13}(1-p_i)$, and the expected length of incubation period, given that symptoms do develop, is calculated using the formula

$$\sum_{i=0}^{13} \frac{(i+1)p_i \prod_{j=0}^{i-1}(1-p_j)}{1-\prod_{i=0}^{13}(1-p_i)}. \quad (17)$$

TABLE I
VARIABLE AND PARAMETERS FOR OAHU MODEL

Parameter, meaning	Value
β, basal transmission rates	optimized to fit data
Factors modifying transmission rate	
ε, asymptomatic transmission	0.75
ρ, reduced healthcare worker interactions	0.8
γ, quarantine	0.2
κ, hospital precautions	0.5
η, healthcare worker precautions	0.5
Population fractions	
p_i, i =0,...,13, onset of symptoms after day i	0.000792, 0.00198, 0.1056, 0.198, 0.2376, 0.0858, 0.0528, 0.0462, 0.0396, 0.0264, 0.0198, 0.0198, 0.0198, 0
$q_{a,i}$, i =0,...,13, asymptomatic quarantine after day i	0 before June 10, then $q_5 = q_6 = q_7 = 0.05$
$q_{s,i}$, i =0,...,4, symptomatic quarantine after day/stage i	C: 0.1, 0.4, 0.8, 0.9, 0.99; H: 0.2, 0.5, 0.9, 0.98, 0.99
r, transition to next symptomatic day/stage	0.2
ν, symptomatic hospitalization	0.08

As we mentioned earlier in the paper, the choice of $q_{a,i}$ reflects the testing and contact tracing efforts in Hawaii. The values of $q_{s,i}$ reflect the sentiment that symptomatic individuals are likely to quarantine, especially after a couple of days of symptoms.

Similarly, parameter r can be viewed as the probability of transitioning from one stage of the illness to the next (with the final stage being recovery or death). Taking into account that the reported length of illness ranges from 2 weeks (for mild cases) to 6 weeks (for severe cases), and that there is a high prevalence of mild and moderate cases, the value of r is chosen to yield an expected length of illness of 17 days. It is calculated using the following formula:

$$2 + \sum_{n=3}^{\infty} \frac{n(n-1)(n-2)}{2} r^3 (1-r)^{n-3} = 2 + \frac{3}{r}. \quad (18)$$

a) Initial Conditions.: The initial values of most variables are zero. The only non-zero values are the number of susceptible individuals in the general community and the healthcare worker community, $S_c(0) = 937711$, $S_h(0) = 15000$, as well as a single not quarantined symptomatic individual, $I_{c,0}(0) = 1$.

III. MODEL FIT AND FORECASTING

Our simulations demonstrate the critical impact of testing and timely contact tracing with adequate quarantine facilities on the number of hospitalizations and required ICU beds. This information is vitally important to immediate disease mitigation strategy.

We also highlight in our work that individual behavior is vital to control the spread of the virus; if we rely solely on testing/contact tracing and quarantine facilities, we would need an unrealistically high success rate to overcome the transmission rate.

A. Fitting the curve from March 6 to August 31, 2020

In this section, we summarize specifics of the model and parameters necessary for an accurate data fit of Oahu data from March 6th to August 27. We use data from the Hawaii Data Collaborative [21] for the count of daily cases as well as active hospitalisations and active ICU beds. The basal rate of transmission β of SARS-CoV-2 in our model is adjusted in time to reflect non-pharmaceutical measures taken by state of Hawai'i during this pandemic. On March 25, the governor imposed a stay-at-home order which was then progressively

Copyright (c) IARIA, 2020. ISBN: 978-1-61208-817-4

lifted throughout the months of May and June. The exact timeline of events can be found here [22], but we focus on five major events as transition points for adjusting the β. Due to an alarming spike in daily cases, U.S. Surgeon General Jerome Adams visited Hawaii for a couple of days starting on August 25 to address the situation with Hawaii government and talk to the public. On August 27, a new stay-at-home order was imposed. This second lockdown is complemented with improved access to testing with a goal of improving contact tracing and to once again stop the spread.

The primary goal of our work is to fit model to current data by selecting appropriate parameter values and then to employ the model in a predictive fashion to improve the mitigation strategy. In general, fitting parameters in an SEIR model is often a combination of mathematics and sociology, since changes in the transmission rate generally occur with a certain time lag based on governmental restrictions and population response [23]. Except for the basal transmission rate, our model parameters are fixed to correspond to available information about the virus and the disease. The basal transmission rates are obtained by optimizing the fit to the data using the Levenberg–Marquardt algorithm and are shown in Table II. A more detailed description of our complete model and parameters is given in Section II.

TABLE II
OPTIMIZED TRANSMISSION RATES TO FIT OAHU DATA. THEY REFLECT THE STATE AND OAHU NON-PHARMACEUTICAL MITIGATION MEASURES.

Transmission rates		
March 6 - April 2	April 2 - May 20	May 20 - May 30
$\beta = 0.3657$	$\beta = 0.0491$	$\beta = 0.1133$
May 30 - June 10	June 10 - Aug 11	Aug 11 - Aug 27
$\beta = 0.2109$	$\beta = 0.1694$	$\beta = 0.1086$

Figure 1 displays the model, run with the transmission rates from Table II, versus the real data. The dots are the recorded daily case counts and the solid line gives the model daily counts.

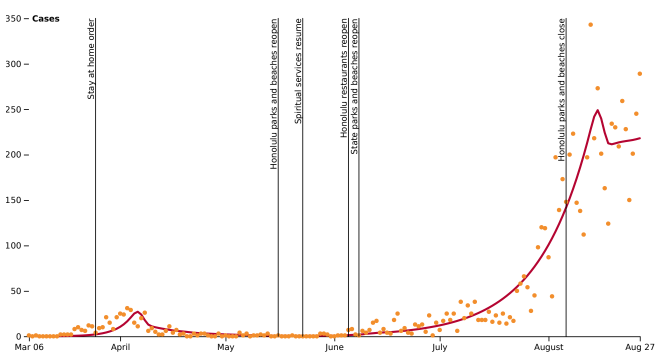

Fig. 1. Daily cases. Dots are the actual data and the plain line represents the model. We also delineate the various mitigation measures that took place during that period.

An important quantifier in COVID-19 is the number of hospitalization and ICU beds. This is a concrete number (as opposed, say, to unknown carriers) that is historically used to document the amount of disease. Additionally, since we have seen hospitals throughout the world being overwhelmed by the number of COVID-19 patients, it is a critical element of mitigation strategy. Figure 2 shows our model values for active hospitalizations (left) and active ICU beds (right) along with the corresponding recorded values. Note that the model values are slightly smaller than the best data fit would yield, however this results from the fact that hospitalisation and ICU data are available for the entire State only, while the model is fitted to the daily cases on Oahu. Also, the data are shown starting July 18, since the numbers for earlier dates have not been released by the Department of Health.

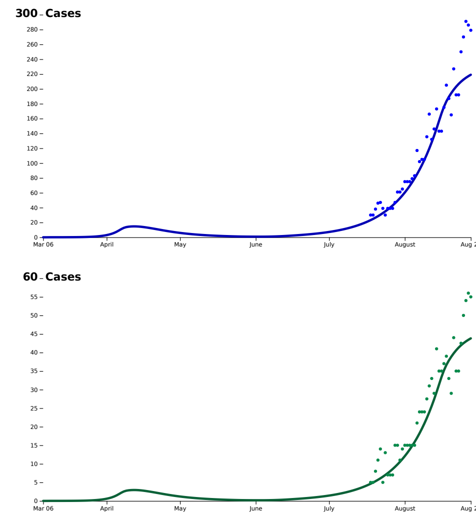

Fig. 2. Data fit for data on active hospitalization (blue) and ICU beds (green). Real data are dots and model predictions are lines.

B. Impact of testing, contact tracing, and quarantine measures

1) Revisiting the Past: In this section, we retrospectively predict the impact on the number of hospitalisations and ICU beds if proper testing/contact tracing and quarantine measures would have been in place on June 10, corresponding to the date when many of the Hawai'i stay-at-home restrictions were lifted.

Some level of testing and contact tracing were in place, but precise data have not been shared publicly and it has turned into a very controversial issue for the State that we will not discuss here. We therefore assume minimal contact tracing and limited testing which we believe is the most accurate representative of the situation. In Figures 1 and 2, we assume that starting June 10, 15% of the asymptomatic people are going into quarantine as the result of testing and contact tracing. More precisely, we assume we catch about 14.3% of asymptomatic population as follows: 5% after day 5 of being exposed, then 5% of the remaining after day 6 of exposure, and then another 5% of the remaining after day 7. We will denote this scenario as $5 : 0.05, 6 : 0.05, 7 : 0.05$ (days 5,6 and 7, each at 5%).

There are several factors which affect the number of asymptomatic individuals going into quarantine, thus slowing down the spread of the virus: improved testing with more rapid turn around, increased contact tracing, and dedicated quarantine facilities.

a) Impact of early asymptomatic quarantine.: Faster detection of asymptomatic individuals can be achieved through a more rapid turnaround on testing as well as through increased contact tracing. Our model parameters reflect this combined effect. Table III shows the impact of the earlier detection on the total number of cases from June 10 to August 27 as well as on the cumulative number of active hospitalisations and active ICU patients for the two and a half month period. These cumulative numbers are computed by summing up the number of all hospitalized (ICU) patients for each day. Table IV shows the same comparison but with a higher percentage of detected asymptomatic individuals going into quarantine.

The financial impact of these measures are addressed in more details in Section IV but just improving time to trace asymptomatic from days 5,6,7 to days 2,3,4 under this minimal contact tracing scenario would save about \$2.6 million.

TABLE III
IMPACT OF DELAY IN ASYMPTOMATIC DETECTION ON THE SPREAD OF COVID-19

Testing/Contact Tracing	Total Cases	Cum act Hospt.	Cum act ICU
5:0.05, 6:0.05, 7:0.05	6517	4721	944
3:0.05, 4:0.05, 5:0.05	5658	4163	833
2:0.05, 3:0.05, 4:0.05	5102	3799	760

Note: Cum Act Hospt. (Cum act ICU) refers to the cumulative number of active hospitalizations (ICU patients).

TABLE IV
IMPACT OF DELAY IN ASYMPTOMATIC DETECTION ON THE SPREAD OF COVID-19

Testing/Contact Tracing	Total Cases	Cum act Hospt.	Cum act ICU
5:0.1, 6:0.1, 7:0.1	5760	3953	791
3:0.1, 4:0.1, 5:0.1	4346	3088	618
2:0.1, 3:0.1, 4:0.1	3551	2590	518

b) Impact of the volume of asymptomatic quarantine.: The actual percentage of detected asymptomatic individuals is affected by the amount of testing done, by the amount of contact tracing resources available, and in large part, by quarantine facilities. Quarantine facilities are particularly important for the Oahu modeling, since a large number of residents live in multi-generational and non-family member shared households. Table V shows how various fractions of quarantined asymptomatic individuals affect the total number of cases, hospitalizations, and ICU occupancy, while the dates of quarantine remain the same. Note that the quarantine fraction of 0.1 on each of the three days leads to the overall 27% detection of asymptomatic cases, 0.2 reaches 48.8%, and 0.3 reaches 65%.

TABLE V
IMPACT OF VOLUME OF ASYMPTOMATIC DETECTION ON THE SPREAD OF COVID-19

Testing/Contact Tracing	Total Cases	Cum act Hospt.	Cum act ICU
5:0.05, 6:0.05, 7:0.05	6517	4721	944
5:0.1, 6:0.1, 7:0.1	5760	3953	791
5:0.2, 6:0.2, 7:0.2	4499	2865	573
5:0.3, 6:0.3, 7:0.3	3573	2175	435

Together, Tables IV and V show that both the detection day of asymptomatic individuals as well as the volume of quarantined asymptomatic individuals have a significant impact on the hospitalisation and ICU occupancy. Our model suggests a larger benefit when asymptomatic individuals are caught early. Combining both of the above factors, we create various scenarios to predict how the total hospitalisation and ICU beds would have been affected. Figure 3 compares the fit based on real data with two scenarios where testing and contact tracing is assumed to be $3 : 0.15, 4 : 0.2, 5 : 0.1$ and $2 : 0.15, 3 : 0.3, 4 : 0.2$, which represents earlier detection of exposed individuals and in larger quantities. While the two scenarios yield continuously decreasing numbers of daily cases, with the second scenario achieving single digits by the end of the year, this does not represent a proper forecasting due to unrealistic assumptions. We address the forecasting question in detail in the next section.

Notice that the gain from this last scenario compared to the control simulation would amount to a savings of about \$10 million.

2) Forecasting: The data fitting and parameter matching specific to our Oahu data allows us to better understand the effects of the various parameters as well as the transmission rate fits. We then use this to provide forecasting scenarios that are dependant on testing/contact tracing and quarantine measures. These forecasting scenarios are critical to guiding Hawaii decision makers.

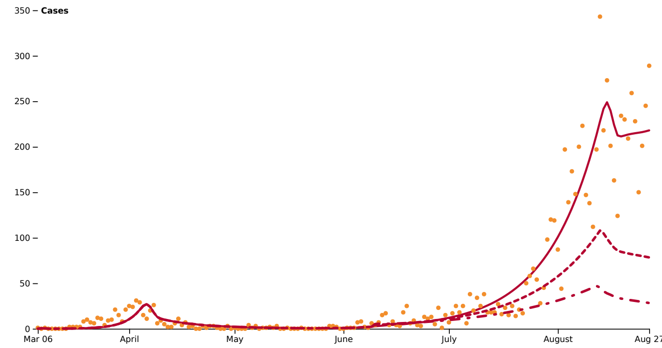

Fig. 3. Comparison of the daily cases real data fit (plain line) to two alternate scenario assuming $3 : 0.15, 4 : 0.2, 5 : 0.1$ (dash line) and $2 : 0.15, 3 : 0.3, 4 : 0.2$ (dotdash line) respectively for testing/contact tracing and quarantine.

TABLE VI
HOSPITALISATION AND ICU VARIATIONS FOR DIFFERENT SCENARIOS

Testing/Contact Tracing	Total Cases	Cum act Hospt.	Cum act ICU
5:0.05, 6:0.05, 7:0.05	6517	4721	944
3:0.15, 4:0.2, 5:0.1	3249	2269	454
2:0.15, 3:0.3, 4:0.2	1667	1208	242

All of our predictive scenarios start on Aug 27 using a transmission rate of $\beta = 0.1086$ from Aug 27 to Aug 30. The transmission rate is then adjusted for each scenario depending on various societal events: stay-at-home order (we assume β slightly higher than during the first stay-at-home order due to community spread); Labor day holiday weekend (increase in transmission rate for a few days); lifting the stay-at-home order on October 5 (varies depending on population behavior), Thanksgiving holiday (variable spike over a few days). At the same time, parameters for testing and contact tracing are gradually adjusted starting August 30 to attain the target values by October 5.

a) Scenario 1.: The first scenario assumes very aggressive testing/contact tracing and facility quarantine but moderate compliance in individual behavior. The target parameter values for testing and contact tracing are taken as $2 : 0.4, 3 : 0.4, 4 : 0.4$, which assumes catching a total of 78% of asymptomatic individuals between days 2 and 4 of exposure. Table VII provides the transmission rates for the various periods. In this scenario, we assume the population will behave similarly to what happened after June 10 once the stay-at-home order is lifted.

TABLE VII
TRANSMISSION RATES FOR SCENARIO 1

Transmission rates		
Aug 30 - Sep 11 $\beta = 0.09$	Sep 11 - Sep 14 $\beta = 0.12$	Sep 14 - Oct 5 $\beta = 0.09$
Oct 5 - Dec 1 $\beta = 0.17$	Dec 1 - Dec 5 $\beta = 0.2$	Dec 5 - Dec 31 $\beta = 0.17$

b) Scenario 2.: The second scenario assumes a more realistic testing/contact tracing and facility quarantine but higher compliance in individual behavior starting after lifting the stay-at-home order on October 5. The target parameter values for testing and contact tracing are taken as $2 : 0.2, 3 : 0.2, 4 : 0.2$, which assumes catching a total of

Copyright (c) IARIA, 2020. ISBN: 978-1-61208-817-4

49% of asymptomatic individuals between days 2 and 4 of exposure. Table VIII provides the transmission rates for the various periods. We assume the population will behave in a more compliant way than what happened after June 10 once the stay-at-home order is lifted. The transmission rate is thus reduced from 0.1694 to 0.145.

TABLE VIII
TRANSMISSION RATES FOR SCENARIO 2

Transmission rates		
Aug 30 - Sep 11 $\beta = 0.09$	Sep 11 - Sep 14 $\beta = 0.12$	Sep 14 - Oct 5 $\beta = 0.09$
Oct 5 - Dec 1 $\beta = 0.145$	Dec 1 - Dec 5 $\beta = 0.2$	Dec 5 - Dec 31 $\beta = 0.145$

c) Scenario 3: The third scenario is identical to scenario 2 but with more a relaxed testing/contact tracing and facility quarantine, reflected by the target testing/contact tracing parameter values of $3 : 0.2, 4 : 0.2, 5 : 0.2$, which still assumes catching a total of 49% of asymptomatic individuals, but now between day 3 and 5 of exposure.

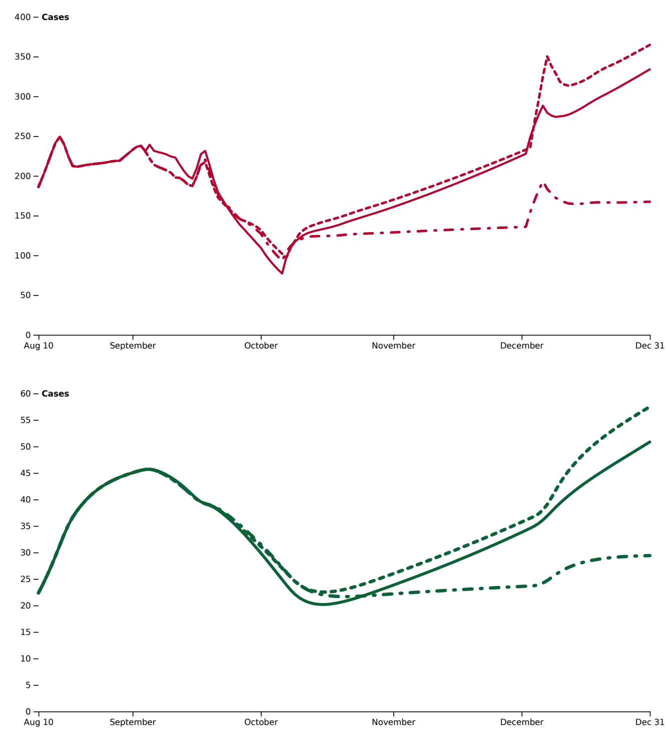

Fig. 4. In red we have the daily cases, and green ICU beds. Scenario 1: plain line. Scenario 2: dot-dash line. Scenario 3: dash line. Scenario 1 is better at first but scenario 2 is provides the best outcome over the long run.

TABLE IX
HOSPITALISATION AND ICU VARIATIONS AS OF DECEMBER 31

Testing/Contact Tracing	Total Cases	Cum act Hospt.	Cum act ICU
Scenario 1	25141	21043	4209
Scenario 2	19318	18162	3632
Scenario 3	26471	22434	4487

In Figure 4 we compare the three scenarios. The plain line corresponds to scenario 1, the dot-dash line is scenario 2 and the dash one is scenario 3. The number of active hospitalisation cases is qualitatively identical to ICU beds, with the first spike around 228 and the dashed line reaching about 287 on December 31. It is important to note that the line for scenario 2 starts to decrease in early 2021, while the number of daily cases for scenarios 1 and 3 see keeps increasing, with a peak of 594 daily cases on April 3 for scenario 1, and a peak of 541 daily cases on April 23 for scenario 3.

IV. SIMULATIONS ANALYSIS

A zoom on Figure 1 for dates between March 6 and May 30 demonstrates the efficiency and good timing of the first stay-at-home order, Hawai'i even being referred at the time as the safest state. Starting in mid-June we see the daily cases increasing and following an exponential trend for a 40 day period to become one of the worst states in dealing with the pandemic. Our simulations attempt to provide an explanation to this observation and produce some forecasting scenarios to help decision makers as we will come out of the second stay-at-home order around October 5.

In Section III-B, we show that with an increased structure of testing/contact tracing and quarantine facilities, we could have dramatically impacted the outcome as of August 27. Our results show that earlier detection of asymptomatic individuals has the most effect on the behavior of the model. From Figure 3 and Table VI, assuming we traced and quarantine successfully 52% of the asymptomatic population after days 2, 3 and 4 (more dominantly after day 3 of being exposed), we would have seen a reduction of 4850 total daily cases, 3513 cumulative active hospitalisation and 702 cumulative active ICU beds which is equivalent to a reduction of about 74% for total daily case, and for both hospitalisation and ICU beds. We can only speculate why these measures were not in place back in June, it was likely due to a combination of different challenges that have been well addressed in the newspapers. For instance, per an article in the Star advertiser on July 8, one of Hawaii's largest COVID-19 testing laboratories supply of chemicals needed to run test locally was restricted and reduced testing capabilities by about 70%. This was a consequence of surge in coronavirus cases across the country and supplies were distributed in priority to states where the intensive care units were overrun. Lack of contact tracing and quarantine facilities for people testing positive have also been issues that prevented the State to successfully keep the count under control. It is easy to downplay quarantine facilities compared to testing and contact tracing, however in Hawaii's a disproportionate fraction of individuals affected by the spread of the virus are Pacific Islanders and they very often live in multi-generational housing that does not permit for isolation. Recently the State has made agreements with hotels to turn them into quarantine facilities.

The forecasting portion of Section III-B raises a very important point. Figure 4 and Table IX demonstrate how different transmission rates and testing/contact tracing, quarantine facilities affect the future of the curve. The take away from these results is that to succeed in controlling the curve, we need a combination of aggressive testing/contact tracing, quarantine facilities as well as compliance from individual to keep the transmission rate to lower levels. Scenario 1 assumes almost perfect success in quarantining exposed individuals but transmission rates comparable to what we had after the State lifted the first stay-at-home order. The second scenario assumes better compliance from the population (lower transmission rate β) and aggressive but doable contact tracing; it provides the best outcome. The scenario 3 with same transmission rate as scenario 2 but shifting the contact tracing by one day shows significantly more cases. The conclusion is that to control the curve long term we need both: aggressive contact tracing and high compliance from the population. Unless the State can enforce such measures, we will be back into a third stay-at-home order in about a couple from months from lifting the second one. Note that for scenario 2, the maximum daily cases will not exceed 193 and the peak will occur in early December due to an assumed increase in non-compliance during the Thanksgiving

Copyright (c) IARIA, 2020. ISBN: 978-1-61208-817-4

holiday, while for scenario 3 we are looking at 541 cases in early April, and we reach 594 cases in late April for scenario 1. Those results are aligned with other work that has been published for instance in [24] [25].

V. Conclusion and Future Work

The goal of our model is not primarily to predict the single most likely outcome for the COVID-19 outbreak on Oahu, but rather is to compare the benefits and costs of implementing various mitigation strategies. We conclude that, provided contact tracing was in place with quarantine facilities as well as explicit guidance for the public on how to behave and compliance to those, we would be now under 50 daily cases and a second stay-at-home would not have been necessary. The State absolutely needs to be prepared when lifting the second stay-at-home order.

A. Economic Impact

It is possible to associate some numbers to the quantitative study we have done. Indeed, calculating the differences between two scenarios for the total hospitalization and ICU beds, we can assign a cost reduction for Oahu over the period June 10 to August 27. Numbers need, however, to be taken with a grain of salt since we are only providing estimates.

We are assuming the cost of a hospitalization day to be $2,000 and we average the daily cost at the intensive care to $4,200. The last number accounts for the fact that a fraction of patients are in need of ventilators and that the first two days in ICU are more costly than the rest of the stay. As mentioned above in the discussion the best scenario in Table VI reduces the total hospitalisation and ICU beds by 74% which amount to almost $10 million. Contact tracing, as well as quarantine facilities also have a cost, but it will be quite lower. Comparing the forecasting scenario, we obtain that as of December 31, scenario 2 saves more than $12 million compared to scenario 3 and scenario 1 saves almost $4 million compared to scenario 3. Those amounts increase quite dramatically after December 31, 2020.

B. Tourism and Vaccines

The State of Hawai'i is, since March 26, 2020, in an effective isolation bubble following the mandatory 14-day traveler quarantine that has not yet been lifted. The interisland quarantine was lifted on June 16 and then partially reinstated on August 11. This is the reason why travelers are not explicitly included in our work; they are currently virtually nonexistent (counts dropped to the lower hundreds from a historical norm of about 30,000 a day). Once tourism is open again, it will be added to the model and new simulations can take place. The form of re-instating tourism is still undetermined, and we expect it will be gradual and measured. Furthermore, if and when a vaccine is developed, our compartmental model can be used to account for the additional sub-population of the vaccinated.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No. 2030789. We would like to thank Prof. Andrea Bertozzi for her insightful comments and feedback.

References

[1] D. Herman, *Shutting Down Hawai'i: A Historical Perspective on Epidemics in the Islands*, 2020, https://www.smithsonianmag.com/history/shutting-down-hawaii-historical-perspective-epidemics-islands-180974506/, retrieved: October, 2020.

[2] E. Blakemore, *How Measles Helped Destroy the Hawaiian Monarchy*, 2019, https://www.history.com/news/hawaii-monarchy-downfall-measles-outbreak/, retrieved: October, 2020.

[3] R. Schmitt and E. Nordyke, "Death in Hawai'i: the epidemics of 1848 - 1849," *The Hawaii Journal of History*, vol. 35, pp. 1–13, 2001.

[4] *The COVID Tracking Project*, 2020, https://www.covidtracking.com/race, retrieved: October, 2020.

[5] M. Webb Hooper, A. M. Nápoles, and E. J. Pérez-Stable, "COVID-19 and Racial/Ethnic Disparities," *JAMA*, vol. 323, no. 24, pp. 2466–2467, 06 2020.

[6] *Covid in the U.S.: Latest Map and Case Count - The New York Times*, 2020, https://www.nytimes.com/interactive/2020/us/coronavirus-us-cases.html, retrieved: October, 2020.

[7] A. Bertozzi, E. Franco, G. Mohler, M. B. Short, and D. Sledge, "The challenges of modeling and forecasting the spread of covid-19," *Proc. Natl. Acad. Sci.*, vol. 117, no. 29, pp. 16 732–16 738, 2020.

[8] M.-A. Rizoiu, S. Mishra, Q. Kong, M. Carman, and L. Xie, "SIR-Hawkes: Linking epidemic models and Hawkes processes to model diffusions in finite populations," in *Proceedings of the 2018 World Wide Web Conference*, 2018, p. 419–428.

[9] M. Gatto *et al.*, "Spread and dynamics of the covid-19 epidemic in italy: Effects of emergency containment measures," *Proc. Natl. Acad. Sci.*, vol. 117, no. 19, pp. 10 484–10 491, 2020.

[10] D. Duque *et al.*, "Timing social distancing to avert unmanageable COVID-19 hospital surges," *Proc. Natl. Acad. Sci.*, vol. 117, no. 33, pp. 19 873–19 878, 2020.

[11] *Compare Trends in COVID-19 Cases and Deaths in States in the US Reported to CDC*, 2020, https://covid.cdc.gov/covid-data-tracker/#compare-trends_newcases.

[12] R_t *COVID-19*, 2020, https://rt.live/.

[13] J. O. Lloyd-Smith, A. P. Galvani, and W. M. Getz, "Curtailing transmission of severe acute respiratory syndrome within a community and its hospital," *Proc. Biol.*, vol. 270(1578), p. 1979–1989, 2003.

[14] R. J. Dy LF, "A covid-19 infection risk model for frontline health care workers," *Netw Model Anal Health Inform Bioinform*, vol. 9(1):57, 2020.

[15] N. J. e. a. Ballard M, Bancroft E, "Prioritising the role of community health workers in the covid-19 response," *BMJ Glob Health*, vol. 5(6):e002550, 2020.

[16] M. Park, A. R. Cook, J. T. Lim, Y. Sun, and B. L. Dickens, "A systematic review of covid-19 epidemiology based on current evidence," *J. Clin. Med.*, vol. 9, no. 4, p. 967, 2020.

[17] "Prevalence of asymptomatic SARS-CoV-2 infection," *Ann. Intern. Med.*, vol. 173, no. 5, pp. 362–367, 2020.

[18] S. M. Moghadas *et al.*, "The implications of silent transmission for the control of covid-19 outbreaks," *Proc. Natl. Acad. Sci.*, vol. 117, no. 30, pp. 17 513–17 515, 2020.

[19] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *J. Soc. Ind. Appl. Math.*, vol. 11, no. 2, pp. 431–441, 1963.

[20] W. M. Getz and E. R. Dougherty, "Discrete stochastic analogs of erlang epidemic models," *J. Biol. Dyn.*, vol. 12, no. 1, pp. 16–38, 2018.

[21] *Hawaii Data Collaborative*, 2020, https://www.hawaiidata.org/covid19, retrieved: October, 2020.

[22] C. Young, *A COVID-19 Timeline: How Honolulu Got To This Point*, 2020, http://www.honolulumagazine.com/Honolulu-Magazine/August-2020/A-COVID-19-Timeline-How-Honolulu-Got-To-This-Point/, retrieved: October, 2020.

[23] G. Chowell, "Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecasts," *Infectious Disease Modelling*, vol. 2, no. 3, pp. 379 – 398, 2017.

[24] D. Kennedy, G. Zambrano, Y. Wang, and O. Neto, "Modeling the effects of intervention strategies on COVID-19 transmission dynamics," *J Clin Virol*, 2020.

[25] M. A. Acuña-Zegarra, M. Santana-Cibrian, and J. X. Velasco-Hernandez, "Modeling behavioral change and COVID-19 containment in mexico: A trade-off between lockdown and compliance," *Math. Biosci.*, vol. 325, p. 108370, 2020.