

DATA ETHICS AND DIGITAL  
PRIVACY IN LEARNING HEALTH  
SYSTEMS FOR PALLIATIVE  
MEDICINE

# STUDIES IN MEDIA AND COMMUNICATIONS

Series Editors: Laura Robinson, Shelia R. Cotten and  
Jeremy Schulz

Volumes 8–10: Laura Robinson and Shelia R. Cotten

Volume 11 Onwards: Laura Robinson, Shelia R. Cotten and Jeremy Schulz

## Recent Volumes:

- Volume 13: *Brazil: Media from the Country of the Future* – Edited by Laura Robinson, Jeremy Schulz and Apryl Williams; Guest Volume Editors: Pedro Aguiar, John Baldwin, Antonio C. La Pastina, Monica Martinez, Sonia Virginia Moreira, Heloisa Pait and Joseph D. Straubhaar; Volume Guest Associate and Assistant Editors: Sayonara Leal and Nicole Speciale
- Volume 14: *Social Movements and Media* – Edited by Jennifer Earl and Deana A. Rohlinger
- Volume 15: *e-Health: Current Evidence, Promises, Perils and Future Directions* – Edited by Timothy M. Hale, Wen-Ying Sylvia Chou and Shelia R. Cotten; Assistant Editor: Aneka Khilnani
- Volume 16: *Media and Power in International Contexts: Perspectives on Agency and Identity* – Edited by Apryl Williams & Laura Robinson; Guest Editor: Ruth Tsuria; Associate Editor: Aneka Khilnani
- Volume 17: *Networks, Hacking and Media – CITAMS@30: Now and Then and Tomorrow* – Edited by Barry Wellman, Laura Robinson, Casey Brienza, Wenhong Chen and Shelia R. Cotten; Associate Editor: Aneka Khilnani
- Volume 18: *The M in CITAMS@30: Media Sociology* – Edited by Casey Brienza, Laura Robinson, Barry Wellman, Shelia R. Cotten and Wenhong Chen
- Volume 19: *Mediated Millennials* – Edited by Jeremy Schulz, Laura Robinson, Aneka Khilnani, John Baldwin, Heloisa Pait, Apryl A. Williams, Jenny Davis, and Gabe Ignatow
- Volume 20: *Theorizing Criminality and Policing in the Digital Media Age* – Edited by Jeremy Schulz and Laura Robinson, and Julie. B Wiest
- Volume 21: *Mass Mediated Representations of Crime and Criminality* – Edited by Jeremy Schulz and Laura Robinson, and Julie. B Wiest
- Volume 22: *Media, Development and Democracy* – Edited by Heloisa Pait and Juliana Laet

# EDITORIAL BOARD MEMBERS

Rebecca Adams  
*University of North Carolina  
Greensboro*

Ron Anderson  
*University of Minnesota*

Denise Anthony  
*University of Michigan*

Alejandro Artopoulos  
*University of San Andrés*

Jason Beech  
*University of San Andrés*

Grant Blank  
*University of Oxford*

Geoffrey C. Bowker  
*University of California, Irvine*

Casey Brienza  
*Media Sociology Preconference*

Jonathan Bright  
*University of Oxford*

Manuel Castells  
*University of Southern California*

Mary Chayko  
*Rutgers University*

Wenhong Chen  
*University of Texas at Austin*

Lynn Schofield  
*Clark University of Denver*

Jenny L. Davis  
*Australian National University*

Hopeton S. Dunn  
*University of the West Indies*

Jennifer Earl  
*University of Arizona*

Joshua Gamson  
*University of San Francisco*

Hernan Galperin  
*University of Southern  
California*

Blanca Gordo  
*International Computer Science  
Institute*

Tim Hale  
*University of Illinois at Urbana-  
Champaign*

David Halle  
*University of California,  
Los Angeles*

Caroline Haythornthwaite  
*Syracuse University*

Anne Holohan  
*Trinity College*

Heather Horst  
*University of Sydney*

Gabe Ignatow  
*University of North Texas*

Samantha Nogueira  
*Joyce Saint Mary's College of  
California*

Vikki Katz  
*Rutgers University*

Nalini Kotamraju  
*Salesforce*

Antonio C. La Pastina  
*Texas A&M University*

Robert LaRose  
*Michigan State University*

Sayonara Leal  
*University of Brasilia*

Brian Loader  
*University of York*

Monica Martinez  
*University of Sorocaba*

Noah McClain  
*Illinois Institute of Technology*

Gustavo Mesch  
*University of Haifa*

Sonia Virginia Moreira  
*Rio de Janeiro State University*

Gina Neff  
*University of Oxford*

Christena Nippert-Eng  
*Indiana University*

Hiroshi Ono  
*Hitotsubashi University*

C. J. Pascoe  
*University of Oregon*

Trevor Pinch  
*Cornell University*

Anabel Quan-Haase  
*University of Western Ontario*

Kelly Quinn  
*University of Illinois at Chicago*

Violaine Roussel  
*University of Paris*

Saskia Sassen  
*Columbia University*

Sara Schoonmaker  
*University of Redlands*

Markus S. Schulz  
*International Sociological  
Association*

Joseph D. Straubhaar  
*University of Texas at  
Austin*

Mike Stern  
*Michigan State University*

Simone Tosoni  
*Catholic University of Milan*

Zeynep Tufekci  
*University of North Carolina, Chapel  
Hill*

Eduardo Villanueva  
*Pontifical Catholic University of  
Peru*

Keith Warner  
*Santa Clara University*

Barry Wellman  
*Ryerson University*

Jim Witte  
*George Mason University*

Simeon Yates  
*University of Liverpool*

STUDIES IN MEDIA AND  
COMMUNICATIONS VOLUME 23

# DATA ETHICS AND DIGITAL PRIVACY IN LEARNING HEALTH SYSTEMS FOR PALLIATIVE MEDICINE

EDITED BY

**VIRGINIA M. MIORI**

*Saint Joseph's University, USA*

**DANIEL J. MIORI**

*Erie County Medical Center, USA*

**FLAVIA BURTON**

*Saint Joseph's University, USA*

AND

**CATHERINE G. CARDAMONE**

*Eigen X, USA*



United Kingdom – North America – Japan  
India – Malaysia – China

Emerald Publishing Limited  
Emerald Publishing, Floor 5, Northspring, 21-23 Wellington Street, Leeds LS1 4DL.

First edition 2024

Editorial Matter and Selection © 2024 Virginia M. Miori, Daniel J. Miori, Flavia Burton and Catherine G. Cardamone.

Individual chapters © 2024 The authors.

Published under exclusive licence by Emerald Publishing Limited.

**Reprints and permissions service**

Contact: [www.copyright.com](http://www.copyright.com)

No part of this book may be reproduced, stored in a retrieval system, transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without either the prior written permission of the publisher or a licence permitting restricted copying issued in the UK by The Copyright Licensing Agency and in the USA by The Copyright Clearance Center. Any opinions expressed in the chapters are those of the authors. Whilst Emerald makes every effort to ensure the quality and accuracy of its content, Emerald makes no representation implied or otherwise, as to the chapters' suitability and application and disclaims any warranties, express or implied, to their use.

**British Library Cataloguing in Publication Data**

A catalogue record for this book is available from the British Library

ISBN: 978-1-80262-310-9 (Print)

ISBN: 978-1-80262-309-3 (Online)

ISBN: 978-1-80262-311-6 (Epub)

ISSN: 2050-2060 (Series)



INVESTOR IN PEOPLE

# CONTENTS

<i>List of Figures and Tables</i>	<i>ix</i>
<i>About the Authors</i>	<i>xiii</i>
<i>Acknowledgments</i>	<i>xiv</i>
<b>Chapter 1 Making the Case</b> <i>Daniel J. Miori</i>	<i>1</i>
<b>Chapter 2 Privacy and Learning Health Systems</b> <i>Daniel J. Miori</i>	<i>9</i>
<b>Chapter 3 Shaping the Continuum of Care Through Public Policy and Data</b> <i>Thomas R. Martin</i>	<i>17</i>
<b>Chapter 4 Public Data Sources: Cleaning and Wrangling</b> <i>Virginia M. Miori</i>	<i>27</i>
<b>Chapter 5 Public Data Sources: Sizing the Palliative Population</b> <i>Virginia M. Miori</i>	<i>61</i>
<b>Chapter 6 Private Data Sources, Data Privacy and Data Simulations for Palliative LHS</b> <i>Virginia M. Miori</i>	<i>79</i>
<b>Chapter 7 Synthea Descriptive Analysis</b> <i>Virginia M. Miori</i>	<i>91</i>
<b>Chapter 8 Palliative LHS Analysis</b> <i>Virginia M. Miori</i>	<i>111</i>
<b>Chapter 9 Data Repository Design for Public Data Analysis</b> <i>Brian W. Segulin</i>	<i>125</i>

<b>Chapter 10 Palliative LHS Development and API to Ensure Data Privacy</b> <i>Brian W. Segulin</i>	<i>137</i>
<b>Chapter 11 Learning Health Systems Ethics Review</b> <i>Daniel J. Miori</i>	<i>151</i>
<i>Index</i>	<i>157</i>

# LIST OF FIGURES AND TABLES

## Figures

Fig. 5.1.	Percentage of Population That Speaks English Less Than Very Well	65
Fig. 5.2.	Percentage of Population at Levels of Educational Attainment	66
Fig. 5.3.	Household Income as a Percentage of Poverty Level	67
Fig. 5.4.	Percentage of Population by Health Insurance Status	68
Fig. 5.5.	Percentage of Population by Employment Status	69
Fig. 5.6.	General Health Status by Age Group	71
Fig. 5.7.	General Health Status by Hypertension Status	73
Fig. 5.8.	General Health by Heart Issue Status	73
Fig. 5.9.	General Health by Cancer Status	74
Fig. 5.10.	General Health by Kidney Disease Status	74
Fig. 5.11.	General Health by Diabetes Status	75
Fig. 5.12.	General Health by COPD Status	75
Fig. 5.13.	General Health by Metabolic Syndrome Status	76
Fig. 5.14.	Palliative Care Potential by Race and General Health Status	76
Fig. 7.1.	Histogram of the Count of Comorbidities in Synthetic Patients	94
Fig. 7.2.	Histogram of the Transformed Count of Comorbidities in Palliative Patients	95
Fig. 8.1.	Kaplan-Meier Curves for Palliative Population Groups 1–3	119
Fig. 8.2.	Hazard Rates for Palliative Patients in Groups 1–3	119
Fig. 9.1.	Logical Representation of DP Table	126
Fig. 9.2.	Cross-Reference Database Tables	128
Fig. 9.3.	SQL Statement to Populate FIELD_NAMES Table	130
Fig. 9.4.	SQL Statement to Populate FIELD_CROSS_REFERENCE	131
Fig. 9.5.	Python Template for Building URL	132
Fig. 9.6.	ACS_VALUES Table Integration into the Database	133
Fig. 9.7.	SQL Query Generated to Store Values in Database	134
Fig. 9.8.	ACS Data View Definition	134
Fig. 9.9.	ACS Data Query	135
Fig. 10.1.	LHS Logical Architecture	139
Fig. 10.2.	Palliative Consultation Landing Page	146
Fig. 10.3.	Palliative Status Display	147
Fig. 10.4.	Kaplan-Meier Curve Request Page	147
Fig. 10.5.	Generated Kaplan-Meier Curves	149

## Tables

Table 3.1.	Overview of Common Electronic Palliative Care Coordination Systems	22
Table 4.1.	Overall Homeless Population/States with Population in Excess of 10,000	29
Table 4.2.	BRFSS Core Sections 2019 and 2020	33
Table 4.3.	BRFSS Optional Modules Utilized in 2019	34
Table 4.4.	BRFSS Optional Modules Utilized in 2020.	34
Table 4.5.	Chronic Conditions Identified/Calculated with BRFSS.	35
Table A4.1.	Data Profile 02: DP02.	37
Table A4.2.	Data Profile 03: DP03.	44
Table A4.3.	Data Profile 04: DP04.	50
Table A4.4.	Data Profile 05: DP05.	56
Table 5.1.	Patient Characteristics to Determine Quality and Availability of Palliative Care.	64
Table 5.2.	HHS Published Poverty Levels.	67
Table 5.3.	General Health by Age Group as % of Race.	72
Table 5.4.	Size of Potential Palliative Population by Race.	77
Table 6.1.	EHR Primary Characteristics.	81
Table 6.2.	Synthea Output Formats.	83
Table 6.3.	Levels for Education, Employment, and Income Variables.	85
Table 6.4.	Groupings for Race, Sex, and Age Group.	85
Table 6.5.	Racial Groupings.	86
Table 6.6.	Age Group Mapping.	86
Table 6.7.	Sample CDFs for Employment Status.	87
Table 7.1.	Comorbidities to Predict Need for Palliative Care.	92
Table 7.2.	Demographic Breakdown of All Patients (30 and Over).	93
Table 7.3.	Demographic Breakdown All Palliative Patients.	96
Table 7.4.	Demographic Breakdown Deceased Palliative Patients.	96
Table 7.5.	Palliative Population by Sex.	97
Table 7.6.	Palliative Population by Race.	97
Table 7.7.	Palliative Population by Age Group.	97
Table 7.8.	Educational Attainment Completion Percentages by Year.	98
Table 7.9.	Education Percentages for Entire Palliative Population.	98
Table 7.10.	Employment Percentages for Entire Palliative Population.	99
Table 7.11.	Insurance Status Percentages for Entire Palliative Population.	100
Table 7.12.	Income Brackets for Entire Palliative Population.	101
Table 7.13.	Income as Percent of Poverty Level for Entire Palliative Population.	101
Table 7.9a.	Incomplete Secondary Education by Sex-Race-Age Group.	103
Table 7.9b.	Secondary Education by Sex-Race-Age Group.	104
Table 7.9c.	Post-Secondary Education by Sex-Race-Age Group.	104
Table 7.10a.	Unemployed Status by Sex-Race-Age Group.	105
Table 7.10b.	Underemployed Status by Sex-Race-Age Group.	106
Table 7.10c.	Well-Employed/Retired Status by Sex-Race-Age Group.	106

Table 7.11a. Uninsured by Sex-Race-Age Group.	107
Table 7.11b. Underinsured by Sex-Race-Age Group.	108
Table 7.11c. Insured by Sex-Race-Age Group.	108
Table 7.13a. Income 0%–150% of Poverty Level by Sex-Race-Age Group.	109
Table 7.13b. Income 150%–250% of Poverty Level by Sex-Race-Age Group.	110
Table 7.13c. Income >250% of Poverty Level by Sex-Race-Age Group.	110
Table 8.1. Mean Income Assignment for Percent of Poverty Level Calculation.	113
Table 8.2. Group Characteristics for Palliative Care.	114
Table 8.3. Income Level and Insurance Status Cross-Tabulation.	115
Table 8.4. Chi-Square Results for Employment by Income.	116
Table 8.5. Chi-Square Results Summary.	116
Table 8.6. Mean and Median Output from the Kaplan-Meier Curve Generation.	117
Table 8.7. Pairwise Tests on Kaplan-Meier Curves.	118
Table 8.8. Education Levels.	122
Table 8.9. Translation Levels.	122
Table 8.10. Employment Levels.	122
Table 8.11. Income Ranges.	122
Table 8.12. Income Levels.	123
Table 8.13. Insurance Levels.	123
Table 8.14. Age Groups.	123
Table 9.1. DP 2 Meta Data Subset.	127
Table 9.2. DATA_PROFILE Database Table Schema.	128
Table 9.3. FIELD_GROUPS Database Table Schema.	129
Table 9.4. FIELD_NAMES Database Table Schema.	129
Table 9.5. FIELD_CROSS_REFERENCE Database Table Schema.	129
Table 9.6. DATA_PROFILE Records.	130
Table 9.7. Selected Fields DP Field Groups.	130
Table 9.8. Cross-Reference Label Example.	131
Table 9.9. US Census ACS Profile URL Template Placeholders.	132
Table 9.10. ACS_VALUES Database Table Schema.	133
Table 9.11. FIPS_CODES Database Table Schema.	133
Table 9.12. CSV File Contents.	134
Table 9.13. ACS Data Values Extracted from Database Using Unique Labels.	135
Table 10.1. API Routines.	140
Table 10.2. Insurance Status.	141
Table 10.3. Educational Attainment.	141
Table 10.4. Employment Status.	141
Table 10.5. Translation.	142
Table 10.6. Comorbidity List.	142
Table 10.7. Sample Size Calculations.	148

*This page intentionally left blank*

# ABOUT THE AUTHORS

**Thomas R. Martin** began his journey in healthcare working at the Cleveland Clinic Foundation. He has served as Project Lead for numerous IT implementations and mobile app developments and is a former director with Healthcare Information and Management Systems Society (HIMSS), a global, cause-based, not-for-profit organization focused on better health through information technology (IT). He has lived and worked overseas in Australia, traveling extensively in Asia Pacific for business. His research seeks to explore the intersection of health information technology and public policy with a focus on health administration using both quantitative and qualitative methods. Dr Martin is the author of articles, book chapters, and thought leadership pieces on the role of Connected Health technologies, telehealth, and health information exchange to support coordinated care.

**Daniel J. Miori**, MS PA-C, is a physician assistant and author. He works on the Palliative and Supportive Care team and is on the ethics committee at Erie County Medical Center in Buffalo, NY, USA. He began work in palliative medicine and bioethics in 2008 and has been active in education as well as having authored material on ethics in both scientific and popular press. He remains ever optimistic for the courage and good in all of us and for our ability to inhabit the moments of clarity and truth which define us, in whatever form those moments take.

**Virginia M. Miori**, Ph.D., has over 21 years of teaching experience and over 14 years of industry experience in developing and implementing statistical and operations research models in the area of supply chain/logistics. She is active in research in the areas of healthcare informatics, healthcare operations, rehab center effectiveness, scheduling, simulation, supply chain, and predictive analytics. Dr Miori has published over 38 scholarly articles, coauthored one text book and is actively engaged in interdisciplinary research, applying data science to institutional and industry problems.

**Brian W. Segulin** is a software developer focusing on process automation. He has over 35 years of experience designing and developing software solutions for integration, scheduling, and adaptive modeling. He has done work in the process industries including metals, glass, oil and gas, paper, and food and beverage. He specializes in integrating legacy systems with state-of-the-art control solutions, focusing on data security.

# ACKNOWLEDGMENTS

We, the authors, humbly acknowledge the great time and energy our team put into creating this book. Brian W. Segulin's help in coding and data preparation was indispensable as was Thomas R. Martin's work and expertise in health information systems. We are also grateful to Kyle Chalmers, Catherine G. Cardamone, and Gabriel Gil Olavarrieta for their hard work. Most especially, however, we wish to thank Flavia Burton for her critical eye and tireless efforts. We know she will go on to do great things.

In addition to the work of those who contributed, we also wish to thank the families of all those who spent their free time bringing this volume to publishable form. You are our foundation.

# CHAPTER 1

## MAKING THE CASE

Daniel J. Miori

### ABSTRACT

*This chapter introduces Learning Health Systems (LHS) and the impact of data science on such systems. It also examines the necessary properties of data used in LHS and identifies patients who may benefit from a transition to palliative care. Finally, it examines the way LHS can be used to identify racial and social disparities in access to palliative care.*

**Keywords:** Learning Health Systems; palliative care; decision support; electronic health record; data science; unconscious bias

“Very well. Sea,” cried Canute, “I command you to come no further! Waves, stop your rolling! Surf, stop your pounding! Do not dare touch my feet.” ([The Changing Story of Cnut and the Waves, 2020](#))

The anticipation that LHSs in healthcare will advance by their own substantial inertia is a central premise to this work. Just as the sea tide did not cease for King Canute, the data analytics tide will not cease for us. LHSs will advance because of the benefits they can provide, and by doing so will cause healthcare to evolve despite the historically sluggish pace of adoption of analytics techniques. As a result of this inexorable advance, discussion and understanding of the appropriate data and analytics techniques along with their ethical implications require attention and review, particularly within the context of the six Vs of big data: volume, variety, velocity, veracity, validity, and volatility.

---

Data Ethics and Digital Privacy in Learning Health Systems for Palliative Medicine  
Studies in Media and Communications, Volume 23, 1–8

Copyright © 2024 by Daniel J. Miori

Published under exclusive licence by Emerald Publishing Limited

ISSN: 2050-2060/doi:[10.1108/S2050-206020230000023001](https://doi.org/10.1108/S2050-206020230000023001)

LHSs can reasonably be seen as having a central role in the overall advancement of healthcare. An LHS is defined by the [Agency for Healthcare Research and Quality \(AHRQ, 2020\)](#) as:

a health system in which internal data and experience are systematically integrated with external evidence, and that knowledge is put into practice. As a result, patients get higher quality, safer, more efficient care, and health care delivery organizations become better places to work.

In the realm of general medicine, an LHS can have a broad focus and still be effective in serving clinicians and patients. In the face of more complicated and chronic conditions however, a broad-based LHS may not serve the purpose of either. As with most large-scale software systems, the creation of an overarching, all-inclusive tool is very difficult. To effectively address all of the intricacies of any health condition, these systems must be narrowly defined and focused on particular patient populations. In this way, an LHS can support medical decision-making. Palliative care is one such area where a narrower definition will improve software system outcomes.

Palliative care is a field with its own philosophical approach and unique challenges which should be considered in discussion of the future of a palliative specific LHS. Palliative care providers often deal with the apprehensions of other clinicians and of patients over a perception of it as “comfort care,” a euphemism for withdrawal of medical treatment followed by the rigorous application of morphine. In reality, palliative care is best characterized as facilitating the transition from futile, potentially burdensome, aggressive, curative type care for a patient with an irreversible (possibly terminal) condition to treatment encompassing goals based on quality of life, as agreed on by practitioner and patient. This approach may not necessarily include a wish to limit attempts at resuscitation as is seen in hospice-type care. Frequently, attaining those quality-based goals may involve some continued highly aggressive care to improve function and limit symptoms.

The ability of clinicians to anticipate, or prognose, future medical events is difficult under ordinary circumstances, and vast research has been done to better understand the potential and medical needs of individual patients. In the United States (US), foundational work such as The Framingham Study ([Yuling Hong, 2020](#)) is an example of long-term large-scale data gathering which has both improved our understanding of a clinical condition (heart disease), and demonstrated the unintended consequences of limited access. The original study group was overwhelmingly white males but it has since been expanded to better represent the American population. In the same way, The Framingham Study improved our understanding of early intervention in the management of heart disease; an LHS can better anticipate the individual patient’s clinical options and the probability of success or failure of specific interventions in establishing and attaining the goals of that patient.

According to the World Health Organization (WHO), only 14% of people (worldwide) who need palliative care, actually receive it ([Palliative Care Fact Sheet, 2020](#)). Though the percentage is much higher in the US, low- and middle-income

individuals have significantly less access to palliative care. The WHO also identified key barriers to palliative care, including:

- “lack of awareness among policy-makers, health professionals and the public about what palliative care is, and the benefits it can offer patients and health systems;
- cultural and social barriers, such as beliefs about death and dying;
- misconceptions about palliative care, such as that it is only for patients with cancer, or for the last weeks of life; and
- misconceptions that improving access to opioid analgesia will lead to increased substance abuse.”

Examining the impact of LHSs through the lens of ethics with regard to data collection and decision-making in palliative care may seem to generate concerns about the role of an LHS due to the fear of the indiscriminate use of a data-dependent prognostic tool in place of a trained clinician in evaluation and discussion of goals of patient care. It is important that an LHS is not be viewed as a decision-making surrogate, but rather as a decision support tool. Unlike artificial intelligence approaches, the heavy lifting of reviewing and implementing potential decisions remains with the clinician and patient. Also, in order for an LHS to assist with decision-making, a reliable and acceptable method for codifying and incorporating prudential and moral norms is a reasonable expectation. The programmers responsible for the logic of such a system would hold ethical (and potentially legal) responsibility for that system, similar to the clinicians who use it.

The intent of this work is to propose the effectiveness of an LHS as a data source to overcome subtle but powerful flaws in the delivery of medical care. In this case, there is equitable access to specialized treatment. Within this paradigm however, other possibilities exist. Ultimately, any LHS must be viewed as a source of additional information for the clinician and patient (or surrogate decision-maker) in much the same way as diagnostic tools such as an MRI or blood tests. Using current industry standards for approving diagnostic devices as a precedent, it becomes easier to suggest that the recommendations of an LHS could be acceptable evidence of the improbability of meeting curative treatment goals within the stated definition of palliative care. This evidence would allow transition to a less burdensome, symptom-control based plan of treatment.

In providing this support, LHS recommendations also have the potential to address issues such as the tendency of clinicians to overestimate prognosis at the end of life. In some circles, overestimation of prognosis can be thought of as kindness, a white lie which makes the final months of a person's life less conflicted and more hopeful. Treating the family by continuing aggressive but futile curative type care is a parallel issue. It has the theoretical purpose of buying time to allow families to gain acceptance of an untimely death and tends to be offered as a rationalization for continuing treatment in the absence of any hope of success. Providing that unrealistic assessment, however, does not help. It simply delays

a potentially pathological grieving process until the family is out of a medical setting and less able to access support services. It seems to be an obvious error, treating emotional trauma by avoidance. Timely referral for grief counseling would be more appropriate, and any delay in seeking such help simply harms the individuals who need it.

An LHS can drive timely and efficient treatment decisions, lowering the overall cost of care, and expanding the availability of palliative care as greater awareness of need across social boundaries occurs. Of course, even the most optimistic assessment of that transition within the US healthcare delivery system must acknowledge that the process will be complex and will generate unforeseeable barriers. Understanding the current landscape of Electronic Health Records (EHR), their strengths, weaknesses, and ongoing advances will be essential in limiting those potential barriers (The terms *Electronic Health Record* and *Electronic Medical Record* have both been used extensively in literature; the distinctions will be addressed in Chapter 6. For our purposes, we will use the term *Electronic Health Record*).

## PALLIATIVE DATA

The nature of healthcare data, and specifically the desired nature of palliative care data, may be seen through six Vs of big data. These include: volume, variety, velocity, veracity, validity, and volatility.

### *Volume*

The amount of data in healthcare is expanding exponentially. Not only are hospitals, doctors, and health care systems increasing their data contribution but also publicly held data relevant to health care delivery is expanding substantially as well. This includes the US Census American Community Survey, the Centers for Disease Control Behavioral Risk Factor Surveillance System, the Bureau of Labor Statistics, the Bureau of Economic Analysis, the Housing and Urban Development (HUD) Exchange, the Migration Policy Institute (MPI), and many other government and nonprofit organizations.

Not only are users faced with thousands of stores of data but also with millions of variables. The data wrangling and cleaning effort to even begin to identify and understand the characteristics of palliative patients is extremely consuming.

### *Variety*

As noted under the section “Volume,” the relevant data sources are extremely varied. Privately held health care data resides in EHR systems with complex relationships and data formats that include free text, audio files, video files, and single- and multi-file DICOM (Digital Imaging and Communications in Medicine) formats, to name just a few. Not only are we faced with types of data when discussing variety, but we are also faced with data reflecting many different facets of a person’s health impacting healthcare decisions.

Publicly available data includes survey responses of all types, geographic location data, nonprofit organization operational data, etc.

### *Velocity*

The velocity of data collection is increasing especially with the advent of the Internet of Things (IoT). Not only is healthcare data collected by physicians' offices, clinics, hospitals, and testing centers, but it is also collected in large volumes by personal medical devices such as heart monitors, glucose monitors, wearables, and myriad smartphone apps. All of this data is important in informing treatment evaluations.

### *Veracity*

Big data veracity references bias and abnormality in data. It is a significant challenge, especially when examining palliative data. In particular, the lack of representation results not only in racial bias but also in overall income and education biases. This area is discussed in detail later in this chapter, within the "Unconscious Bias" section.

### *Validity*

The generation of meaningful analysis is irrevocably tied to the correctness and accuracy of data. It may seem like a clear concept, but accuracy is directly related to the data collection effort and a universal understanding of what accuracy in healthcare data collection truly means. Data collected via surveys is inherently limited in accuracy due to self-reporting bias. EHR data benefits in having trained professionals enter data consistently, though it may still be subject to entry and judgment bias. The ability to collect data through standardized tests, digital processes, and the use of wearables is extremely valuable, as well as consistent. Furthermore, the question of what constitutes complete palliative healthcare data must be addressed.

### *Volatility*

Volatility considers the length of time that data is considered to be valid. Palliative conditions and diagnostic data will not have an effective expiration date, however as treatments evolve, prognostic data and treatment data evolve as well. In addition, as data is collected in more diverse types (as mentioned in Variety), analyses adapt to include vastly more data. Older data, lacking this depth and breadth, becomes obsolete and is limited in the contribution it can make to analysis.

## **UNCONSCIOUS BIAS**

After consideration of these six Vs, any discussion of Data Science integration into an LHS for palliative care must thoroughly address unconscious bias due to lack of representation of racial, ethnic, and socio-economic groups in treatment

data. It is also critical to address the impact of this bias on prognoses and diagnoses rendered for patients. An LHS shares the ethical responsibilities of the clinicians who manage patients and make recommendations for medical treatment and should consider questions such as:

- What are survival rates for attempted cardio-pulmonary resuscitation in palliative patients with specific comorbidities?
- What is the likelihood of post-surgical complications for a cancer survivor with a specific treatment profile?
- What is the likelihood of a patient with compromised immunity developing clostridium difficile colitis (c-diff) during a hospital stay?
- What is the likelihood of an elderly patient developing hospital associated delirium?

In palliative care, as well as the practice of medicine in general, timely decisions often need to be made on incomplete evidence. Time gained by catching the patient earlier in an irreversible disease process allows for more effective clinical treatment to relieve symptoms, and greater benefit from time gained in a therapeutic environment to deal with the emotional trauma of illness and death.

Hospitals and medical practitioners with access to LHSs fed by representative data would have greater ability to provide early and equal access to care.

The diverse sources for publicly available data include the US Census Bureau, the US Department of HUD, the Homeless Research Institute, the MPI, the Refugee Processing Center (RPC), the Centers for Medicare and Medicaid Services (CMS), the Center to Advance Palliative Care (CAPC), the Centers for Disease Control (CDC), the National Health Interview Survey (NHIS) and the WHO, among others.

Age, sex, race, educational attainment, ethnicity, income (poverty status), health insurance status, urban versus rural, homeless status, documentation status for immigration, and refugee status should all be used to characterize the US population data and the palliative population. The palliative patient population tends to have increased representation in higher income categories, higher educational levels, and white non-Hispanic individuals. Representation falls well below expectations with populations that are marginalized based on race, lower income, poor English language acquisition, lower educational attainment, and homelessness.

Finding subtle differences in data can identify individuals who would welcome and benefit from a palliative care approach and allows healthcare professionals, aided but not directed by an LHS, to help correct known biases and determine specific factors and analyses that help correct underrepresentation in the palliative population.

## HEURISTICS IN PALLIATIVE CARE

Heuristics are fundamental to the human ability to analyze the world around us. They are sets of rules, or rules-of-thumb, that can be applied to decision-making

based on a limited set of data. These cognitive short-cuts allow decisions to be made in the absence of full, possibly unobtainable, information. A phenomenon known in psychology as *the illusion of control* is a tendency to overestimate one's ability to control events and is one of the known flaws in medical heuristics. It is often referred to as the therapeutic illusion and can be mitigated with simple conscious strategies to balance unconscious heuristics that influence clinical decision-making.

If we accept that poor decision-making can be due to an overestimate of control and the inclination toward confirmation bias (the tendency to accept information that confirms our existing beliefs), that the use of data science based diagnostic tools to uncover patterns that would be difficult to identify in any other way can be of benefit, and that this will improve the quality of diagnosis and prognosis, the use of an LHS to assist in medical decision-making becomes less daunting.

## **DATA SCIENCE FACILITATING TRANSITION TO PALLIATIVE CARE**

Infomarkers, such as terms including “Death Anxiety” and “Anticipatory grieving,” when used in electronic medical record documentation can trigger consideration for palliative consultation (Yao et al., 2015). This and other approaches support the use of prognostic tools (Greer, Jackson, Meier, & Temel, 2013; Smith et al., 2012; Yao et al., 2015). By adding data science analyses, outcomes and conclusions, more specific and definitive indicators can be added to the decision to transition to palliative care. In particular, tools like Natural Language Processing (NLP), text mining, unstructured data analysis (including DICOM formats), and machine learning among other tools assist in the transition from aggressive curative type care to aggressive symptom management.

## **OTHER DATA SCIENCE CONCERNS IN PALLIATIVE CARE**

When creating a LHS, it is critical to understand and accommodate concerns associated with the use of data analytics or data science. They include but are not limited to control over access as well as use and governance of data; transfer of data from healthcare systems to for-profit systems (such as seen with a release of data from the British National Health Service to Google DeepMind), and transfer of poorly controlled, publicly available, health data to healthcare. This includes information generated by an IP address and not an individual, or directly obtained or eavesdropped data from uncontrolled non-medical biometric monitoring devices through personal fitness trackers.

These concerns may give a medical practitioner pause, but we submit them as a foundation for the breadth of work necessary to develop an LHS with the appropriate data security and maintenance of data privacy.

## DATA SCIENCE AND LHS ADDRESS DISPARITIES IN ACCESS TO PALLIATIVE CARE

Although improved access to palliative care is certainly a worthy goal of any work in data science, the focus of this book will be to identify disparities in access to palliative care across racial and social groups in the US. We will propose a system which will use characteristics found in readily sourced large data sets to identify individuals who may not ordinarily be aware of or receive referral to specialized palliative services.

### REFERENCES

- Agency for Healthcare Research and Quality. (2020, September 28). Retrieved from <https://www.ahrq.gov/>
- Greer, J. A., Jackson, V. A., Meier, D. E., & Temel, J. S. (2013). Early integration of palliative care services with standard oncology care for patients with advanced cancer. *CA: A Cancer Journal for Clinicians*, 63(5), 349–363.
- Palliative Care Fact Sheet. (2020, September 28). Retrieved from <https://www.who.int/news-room/fact-sheets/detail/palliative-care>
- Smith, T. J., Temin, S., Alesi, E. R., Abernathy, A. P., Balboni, T. A., Basch, E. M., ... Von Roenn, J. H. (2012). American Society of clinical oncology provisional clinical opinion: The integration of palliative care into standard oncology care. *Journal of Clinical Oncology*, 30(8), 880–887.
- The Changing Story of Cnut and the Waves. (2020, September 7). Retrieved from <http://www.medievalists.net/2015/05/the-changing-story-of-cnut-and-the-waves/>
- Yao, Y., Stifter, J., O Ezenwa, M., Lodhi, M., Khokhar, A., Ansari, R., ... Wilkie, D. (2015). Infomarkers for transition to goals consistent with palliative care in dying patients. *Palliative & Supportive Care*, 13(5), 1427–1434.
- Yuling Hong, M. D. (2020, September 1). The Framingham Heart Study. Retrieved from <https://www.nhlbi.nih.gov/science/framingham-heart-study-fhs>