

BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: Koleck, Theresa

eRA COMMONS USER NAME (credential, e.g., agency login): TIMCHECKTA

POSITION TITLE: Assistant Professor

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

INSTITUTION AND LOCATION	DEGREE <i>(if applicable)</i>	Start Date MM/YYYY	Completion Date MM/YYYY	FIELD OF STUDY
University of Pittsburgh, Pittsburgh, PA	B.S.N	08/2007	04/2011	Nursing
University of Pittsburgh, Pittsburgh, PA	Ph.D.	05/2011	05/2016	Nursing, Health Care Genetics Minor
Columbia University, New York, NY	Postdoctoral	06/2016	05/2018	Nursing, Biomedical Informatics

A. Personal Statement

My program of research is dedicated to mitigating symptom burden in patients diagnosed with chronic conditions using omics-based approaches and informatics/data science techniques. My genomics-intensive (T32NR009759) doctoral studies at the University of Pittsburgh School of Nursing focused on gaining a greater understanding of the biological foundations of cognitive dysfunction in breast cancer survivors. My National Institute of Nursing Research (F31NR014590) and American Cancer Society (DSCN-14-076-01-SCN) funded dissertation work provided evidence that heterogeneity in the biology of breast cancers is associated with variability in pretreatment cognitive dysfunction in survivors. While my dissertation study on a single symptom was highly informative, symptoms in chronic conditions often do not occur in isolation. Appreciating this, I chose to broaden my work to multiple, co-occurring symptoms (i.e., symptom clusters). Symptom clusters, however, are understudied and poorly understood with the majority of symptom cluster research limited to cross-sectional studies of cancer populations using predetermined symptom inventories. There is a significant need to study and phenotypically characterize symptom clusters across chronic conditions; yet, a lack of access to large, longitudinal data sets (that are required to study these complex phenotypes) has limited progress. Recognizing the great potential in using electronic health records (EHRs) and data science techniques to innovatively address this limitation, I obtained additional training in biomedical informatics and data science as a postdoctoral research fellow at Columbia University School of Nursing (T32NR007969). During my fellowship, I learned how to perform symptom-related EHR clinical data mining using the Columbia University Medical Center Clinical Data Warehouse. Currently (K99/R00 NR017651), I am working to develop, apply, and implement a pipeline for characterization of symptom clusters in chronic conditions using both structured (i.e., diagnosis codes) and unstructured (i.e., clinical notes) EHR data. During the K99-phase of the award as an Associate Research Scientist at Columbia University, I obtained knowledge and skills related to essential data science techniques, including clinical data mining, natural language processing, unsupervised machine learning, and data visualization. For the R00-phase as an Assistant Professor at the University of Pittsburgh, I will use the University of Pittsburgh Medical Center Clinical Data Warehouse, Neptune, to examine the reproducibility of the rigorously developed EHR symptom extraction pipeline and identified symptom clusters in the K99-phase of the award. This work represents the essential first step toward future biomarker discovery and clinical application to alleviate symptom burden. As a highly specialized nurse scientist trained in genomics and biomedical informatics, I am at the forefront of advancing symptom science for chronic conditions through pioneering synergy of symptom clusters, omics, and informatics.

B. Positions and Honors

Positions and Employment

- 2010 – 2011 Nurse Intern/Patient Care Technician, University of Pittsburgh Medical Center Magee-Womens Hospital, Oncology and Surgery, Pittsburgh, PA
- 2011 – 2012 Graduate Student Researcher, University of Pittsburgh Medical Center Magee-Womens Research Institute and Foundation, Pittsburgh, PA
- 2012 – 2012 Summer Genetics Institute, National Institute of Nursing Research, Bethesda, MD
- 2012 – 2016 Graduate Student Researcher, University of Pittsburgh School of Nursing, Pittsburgh, PA
- 2015 – 2015 Summer Institute in Statistical Genetics, University of Washington, Seattle, WA
- 2016 – 2016 Summer Institute in Statistics for Big Data, University of Washington, Seattle, WA
- 2016 – 2018 Postdoctoral Research Fellow, Columbia University School of Nursing, New York, NY
- 2018 – 2018 Summer Institute in Statistics for Big Data, University of Washington, Seattle, WA
- 2019 – 2019 Clinical Data Boot Camp, Case Western Reserve University School of Nursing, Cleveland, OH
- 2018 – Pres Associate Research Scientist, Columbia University School of Nursing, New York, NY

Academic and Professional Honors

- 2019 All of Us Research Program Researcher Workbench Alpha-Phase User Tester
- 2019 Invited Speaker on targeted cancer therapies at the International Society of Nurses in Genetics Regional Meeting, Pittsburgh, PA
- 2018 Featured in Columbia Nursing Magazine “Masters of Information” article:
https://issuu.com/columbianursing/docs/columbia_nursing_fall_2018
- 2018 Invited Speaker on the state of data science in nursing and genomics at the International Society of Nurses in Genetics World Congress, Orlando, FL
- 2016 University of Pittsburgh School of Nursing Outstanding Graduate Research Award
- 2016 Invited Speaker on symptom science at the Excelsa Health 2nd Annual Dissemination of Nursing QI, Evidence, & Research Forum, Westmoreland Hospital, Greensburg, PA
- 2016 Invited Speaker on finding and selecting a postdoctoral position at the Columbia School of Nursing Doctoral Research Seminar Series, Columbia University School of Nursing, New York, NY
- 2016 Invited Speaker on the impact of donations on cancer research for the American Cancer Society to community members, 5 presentations throughout the Greater Pittsburgh, PA area
- 2015 Featured on the International Society of Nurses in Genetics, Meet our Members webpage:
http://www.isong.org/ISONG_members.php
- 2015 University of Washington Summer Institute in Statistical Genetics Scholarship & Travel Award
- 2015 Recipient of Sigma Theta Tau International Eta Chapter Ruth Perkins Kuehn Scholarship
- 2014 Invited Speaker on symptom science at the International Society of Nurses in Genetics World Congress, Scottsdale, AZ
- 2013 Recipient of the University of Pittsburgh School of Nursing Bessie Li Sze Memorial Scholarship
- 2012 Recipient of the University of Pittsburgh School of Nursing Bessie Li Sze Memorial Scholarship

Other Experience and Professional Memberships

- 2016 – Pres Member, American Medical Informatics Association (AMIA)
- 2012 – Pres Member, Sigma Theta Tau International (STTI) Honor Society of Nursing
Research Committee Chair, Eta Chapter (08/2017 – Present)
Leadership Succession Committee Chair, Eta Chapter (05/2015 – 07/2016)
Leadership Succession Committee Member, Eta Chapter (07/2014 – 04/2015)
- 2012 – Pres Member, International Society of Nurses in Genetics (ISONG)
2019 ISONG Congress Planning Committee (01/2019 – 04/2019)
2018 ISONG Congress Planning Committee, Poster Blitz Chair (12/2017-10/2018)
2017 ISONG Congress Planning Committee, Poster Blitz Chair (12/2017 – 11/2018)
2017 ISONG Congress Planning Committee, Poster Blitz Chair (09/2016 – 11/2017)
2016 ISONG Congress Planning Committee, Poster Blitz Chair (12/2015 – 08/2016)
2015 ISONG Congress Planning Committee, Poster Blitz Chair (01/2015 – 11/2015)
Student Task Force Committee Member (02/2015 – 01/2016)
- 2015 – 2016 Member, American Society of Human Genetics (ASHG)
- 2011 – 2016 Member, University of Pittsburgh Doctoral Nursing Student Organization (DNSO)
Secretary (08/2012 – 04/2016)
- 2011 – 2016 Member, Eastern Nursing Research Society (ENRS)
- 2012 – 2014 Member, Oncology Nursing Society (ONS)

C. Contributions to Science

1. I am **using and developing data science techniques to study symptoms documented in electronic health records** through my career development work (K99NR017651) and recently completed postdoctoral fellowship (T32NR007969) at Columbia University School of Nursing. In terms of unstructured data, we conducted a systematic review (a) to synthesize the literature on the use of natural language processing to analyze symptom information documented in electronic health record free-text narratives. We found that the focus of the field is on the development of methods to extract symptoms and use symptom information for disease classification tasks rather than the examination of symptoms themselves. These findings highlight the need for the future natural language processing studies that concentrate on the investigation of symptoms. To fill this gap, Dr. Maxim Topaz and I are collaborating to create a comprehensive vocabulary that can be used to extract symptom terms from the clinical notes using NimbleMiner (a software that allows for the creation of lexicons of similar terms from word embedding models). For structured data (abstract b), we are attempting to improve the usability of diagnosis codes to study symptoms through symptom enrichment via normalization. I have also contributed electronic health record and symptom science case studies to book chapters (c) and (d). The case studies focus on the use of medication information for electronic health record-based phenotyping of postoperative nausea and vomiting follow abdominal gynecological surgeries.
 - a. **Koleck, T.A.**, Dreisbach, C., Bourne, P.E., & Bakken, S. (2019). Natural language processing of symptoms documented in free-text narratives of electronic health records: a systematic review. *Journal of the American Medical Informatics Association*, 26(4), 37-46. PMID: PMC6657282
 - b. **Koleck, T.A.**, Bakken, S., & Tatonetti, N.P. (2019, November). Chronic condition symptom enrichment from electronic health records. Poster presentation at the American Medical Informatics Association 2019 Annual Symposium, Washington, DC.
 - c. Bakken, S., **Koleck, T.A.**, Dreisbach, C., Hickey, K.T. (2020). Precision health from a big data perspective. In S. Dorsey & A. Starkweather (Eds.) *Genomics of Pain and Co-Morbid Symptoms*. New York, NY: Springer Nature.
 - d. Bakken, S. & **Koleck, T.A.** (2019). Big data challenges from a nursing perspective. In M. Househ, A. Kushniruk, & E. Borycki (Eds.) *Big Data Big Challenges: A Healthcare Perspective*. New York, NY: Springer.
2. I have made foundational contributions to **understanding the biological mechanisms of variation in the incidence and severity of cancer- and treatment-related cognitive dysfunction experienced by women diagnosed with breast cancer** through my doctoral work (T32NR009759, F31NR014590) and ongoing collaborations with investigators at the University of Pittsburgh School of Nursing. Specifically, our team found that heterogeneity in the biology of breast cancers is associated with variability in the presence and/or severity of cognitive dysfunction. This line of investigation has the potential to lead to the discovery of biomarkers that will assist healthcare professionals in identifying breast cancer survivors most at risk for cognitive dysfunction and targeting at-risk individuals for early and/or intensive interventions to mitigate these risks.
 - a. Merriman, J., Sereika, S.M., Conley, Y.P., **Koleck, T.A.**, Zhu, Y., Phillips, M.L., Bertocci, M.A., Brufsky, A.M., & Bender, C.M. (2018). Exploratory study of associations between DNA repair and oxidative stress gene polymorphisms and cognitive problems reported by postmenopausal women with and without breast cancer. *Biological Research for Nursing*, 21(1):50-60. PMID: PMC6700884.
 - b. **Koleck, T.A.**, Bender, C.M., Clark, B.Z., Ryan, C.M., Ghotkar, P., Brufsky, A., McAuliffe, P.F., Rastogi, P., Sereika, S.M., & Conley, Y.P. (2017). An exploratory study of host polymorphisms in genes that clinically characterize breast cancer tumors and pretreatment cognitive performance in breast cancer survivors. *Journal of Breast Cancer: Targets and Therapy*, 9, 95-110. PMID: PMC5344452.
 - c. **Koleck, T.A.**, Bender, C.M., Sereika, S.M., Ryan, C.M., Ghotkar, P., Brufsky, A.M., Jankowitz, R.C., McAuliffe, P.F., Clark, B.Z., & Conley, Y.P. (2017). Associations between pathologic tumor features and preadjuvant therapy cognitive performance in women diagnosed with breast cancer. *Cancer Medicine*, 6(2), 339-348. PMID: PMC5313647.
 - d. **Koleck, T.A.**, Bender, C.M., Sereika, S.M., Ahrendt, G., Jankowitz, R.C., McGuire, K.R., Ryan, C. & Conley, Y.P. (2014). Apolipoprotein E genotype and cognitive function in postmenopausal women with early stage breast cancer. *Oncology Nursing Forum*, 41(6), E313-25. PMID: PMC4406250.

3. I have contributed to advancement of symptom science through the **innovative application and development of omics methods to symptom data**. Below are some examples of how I have used and expanded our omics toolbox. In article (a), we developed a novel approach for the identification and prioritization of biologically plausible candidate genes for future investigations of the association between genetic variation and symptoms experienced by breast cancer survivors based on clinically-relevant genes examined in prognostic multigene expression profiles for breast cancer. Article (b) uses the pathway analysis methods described in article (a) to identify biologically plausible pathways and candidate genes for future investigations on the genetic underpinnings of musculoskeletal pain during treatment with aromatase inhibitors in breast cancer survivors. In article (c), we developed a method to calculate weighted genetic risk scores from single nucleotide polymorphisms (SNPs) with previously unknown effects that assigns greater risk or protection to SNPs with stronger associations. This method allowed us to evaluate the collective effect of multiple genetic variants on cognitive performance in breast cancer survivors. Article (d) is a review for nurse scientists of rationale, methodological considerations, and key technologies for various genomic research approaches. I contributed a section on candidate gene association studies.
- Koleck, T.A.** & Conley, Y.P. (2016). Identification of candidate genes for symptom variability in breast cancer survivors based on disease characteristics at the cellular level. *Journal of Breast Cancer: Targets and Therapy*, 8, 29-37. PMID: PMC4790538.
 - Zhu, Y., **Koleck, T.A.**, Bender, C.M., & Conley, Y.P. (2019). Genetic underpinnings of musculoskeletal pain during treatment with aromatase inhibitors for breast cancer: a biological pathway analysis. *Biological Research for Nursing*, Epub ahead of print.
 - Koleck, T.A.**, Bender, C.M., Sereika, S.M., Brufsky, A.M., Lembersky, B.C., McAuliffe, P.F., Puhalla, S.L., Rastogi, P., & Conley, Y.P. (2016). Polymorphisms in DNA repair and oxidative stress genes associated with pre-treatment cognitive function in breast cancer survivors: An exploratory study. *SpringerPlus*, 5, 422. PMID: PMC4826652.
 - Baumgartel, K., Zelazny, J., **Timcheck, T.**,* Snyder, C., Bell, M., & Conley, Y. (2012). Molecular genomic research designs. *Annual Review of Nursing Research*, 29, 1-26. PMID: PMC3422781.
* Please note that I previously published under my maiden name, *Timcheck*.
4. I am also contributing to the **advancement of symptom science in underserved populations** through my work as part of the Precision in Symptom Self-Management (PriSSM) Center (P30NR016587), which has the goal of advancing the science of symptom self-management for Latinos through a social ecological lens, and recently completed Reducing Health Disparities through Informatics (T32NR007969) postdoctoral fellowship at Columbia University School of Nursing. Abstract (a) was presented at the 17th World Congress of Medical and Health Informatics and highlights three area of the PriSSM Center – a Latino data repository, information visualization, and center evaluation. Article (b) addressed multiple gaps about the Latino symptom experience. Using a large, community-based survey of urban Latino adults of primarily Dominican heritage, we examined associations between symptoms – pain, fatigue, depressed mood, anxiety, and sleep disturbance – and participation in usual activities as well as satisfaction with participation in social roles. In article (c), our analysis, which controlled for levels of depression, anxiety, and sleep disturbance, revealed that in Latino adults with hypertension having adequate health literacy was associated with a higher adherence score to antihypertensive medications. Article (d) represents a synthesis of the literature on information visualizations of symptoms included as National Institute of Nursing Research common data elements designed for patient and/or healthcare provider communication. This work is especially important for individuals with limited English proficiency, an understudied, health disparate population.
- Bakken, S., Arcia, A., **Koleck, T.**, Merrill, J.A., & Hickey, K.T. (2019). Informatics and data science for the Precision in Symptom Self-Management Center. *Studies in Health Technology and Informatics*, 264, 1827-1828.
 - Koleck, T.A.**, Suero-Tejeda, N., & Bakken, S. (2018). The influence of Latino symptom experience on participation in usual activities and satisfaction with participation in social roles. *Hispanic Health Care International*, 16(3), 134-144. PMID: PMC6541912.
 - Lor, M., **Koleck, T.A.**, Bakken, S., Yoon, S., & Dunn Navarra, A.M. (2019). Association between health literacy and medication adherence among Hispanics with hypertension. *Journal of Racial and Ethnic Health Disparities*, 6(3), 517-524. PMID: PMC6545226
 - Lor, M., **Koleck, T.A.**, & Bakken, S. (2019). Information visualizations of symptom information for patients and providers: a systematic review. *Journal of the American Medical Informatics Association*, 26(2), 162-171. PMID: PMC6657383.

Complete List of Published Work in MyBibliography

<https://www.ncbi.nlm.nih.gov/sites/myncbi/theresa.koleck.1/bibliography/48581904/public/?sort=date&direction=ascending>

D. Additional Information: Research Support and/or Scholastic Performance

Ongoing Research Support

R00NR017651 Koleck (PI) 06/01/20-04/30/23
Advancing Chronic Condition Symptom Cluster Science Through Use of Electronic Health Records and Data Science Techniques
The overall objective of this study is to develop, apply and refine, and implement an optimized data processing and analysis pipeline for the characterization of symptom clusters in common adult chronic conditions for use with electronic health record data. The long term training goal of this project is to assist Dr. Koleck in becoming an independent investigator conducting a program of research dedicated to mitigating symptom burden in patients with chronic conditions through use of omics and informatics. In the R00 phase of this award, Dr. Koleck will independently implement the pipeline developed during the K99 phase using data from another medical center – the University of Pittsburgh Medical Center – to determine the reproducibility of extraction and preprocessing steps, symptom cluster analysis methods, and identified symptom clusters.
Role: Principal Investigator

Completed Research Support

K99NR017651 Koleck (PI) 06/01/18-03/31/20
Advancing Chronic Condition Symptom Cluster Science Through Use of Electronic Health Records and Data Science Techniques
The overall objective of this study is to develop, apply and refine, and implement an optimized data processing and analysis pipeline for the characterization of symptom clusters in common adult chronic conditions for use with electronic health record data. The long term training goal of this project is to assist Dr. Koleck in becoming an independent investigator conducting a program of research dedicated to mitigating symptom burden in patients with chronic conditions through use of omics and informatics. During the K99 phase of this award, we focused on the creation of rigorous symptom extraction and preprocessing pipeline steps for both structured (i.e., billing code) and unstructured (i.e., clinical narrative) symptom data using EHRs from the Columbia University Medical Center Clinical Data Warehouse.
Role: Principal Investigator

T32NR007969 Bakken (PI, Program Director) 06/01/16-05/31/18
Reducing Health Disparities Through Informatics (RHeaDI)
The ongoing goal of the RHeaDI research training program is to prepare nurse scientists with knowledge and skills in informatics and in dissemination and implementation science to address the scientific challenges and opportunities related to use of health information and communication technologies to reduce health disparities and facilitate evidence-based practice.
Role: Trainee