

Investing in Medically Underserved Communities

Capitalizing on Emerging Informatics to Strengthen Preventive Health Measures

Mara Reisinger

SAT 5424 Population Health Informatics

Spring 2024

Present State of Affairs

- As of 2020, 5.3% of American adults (35+ years) are receiving recommended evidence-based preventive health care. This is down from 8.5% in 2015⁽¹⁾
- 69.6% of American adolescents (12-17 years) had a preventive health care visit in the previous year (2020-2021), down from 78.7% in 2016-2017⁽¹⁾
- Healthy People 2030 has a mission to improve these numbers to 11.5% for adults and 82.6% for adolescents

Impact of HIT⁽⁴⁾

- Early monitoring and detection of disease is easier than ever before with the use of data-mining, machine learning, and big data analytics
- Data is transformed into useful information using methods and technologies provided by data mining
- Cost-effective and efficient prevention methods are made possible due to recent developments in data analytics
- Predictive patterns can be improved using machine learning

Data Mining⁽⁵⁾

- Anomaly Detection
 - Used to discover changes in the data set
- Clustering
 - Describes the data using identified categories
- Classification
 - Data is classified into predefined classes
 - Statistical, decision tree, swarm intelligence, k-nearest neighbor, logistic regression, Bayesian classifier, support vectors

Data Analytics⁽⁶⁾

- **Descriptive analytics**
 - Converts data into information used to understand and analyze healthcare decisions
- **Predictive analytics**
 - Detects patterns of connections by analyzing historical data and uses these to predict future trends
- **Prescriptive analytics**
 - Narrows down health problems using medical knowledge and data
- **Discovery analytics**
 - Discovers new information by analyzing known knowledge

Machine Learning⁽⁷⁾

- Supervised learning
 - Uses previous examples to train prediction and classification algorithms
 - Includes Decision Trees, Support Vector Machines, and Artificial Neural Networks
- Unsupervised learning
 - Uses clustering algorithms to sort data into classes or categories
 - Includes k-Means algorithm, Deep Belief Networks, and Convolutional Neural Networks
- Reinforcement learning
 - Forms strategies based on desired and undesired outcomes
 - Includes Recurrent Neural Network

Challenges

- Data Mining⁽⁵⁾
 - Data sets are imbalanced, incomplete, or too small
 - One model often doesn't work, requiring the use of several to resolve any issues
- Data Analysis⁽⁴⁾
 - Successful analysis relies on the availability, quality, and complexity of data
- Machine Learning⁽⁷⁾
 - Validity related skepticism of the reliance on probability of error, learning relies on the quality of data, and ethical concerns

Current Research in Population Health using Informatics⁽⁸⁾

Health People 2030 is the largest collection of population health data in the United States. The data comes from more than 80 sources that all meet the same meticulous criteria. The data must be nationally representative, publicly available, and have known population coverage, response rates, and documentation completeness. Baseline data must be from more recently than 2015 and able to provide two additional data points throughout the decade.

Data is analyzed and tracked by the National Center for Health Statistics

Opportunities⁽⁴⁾

- Incorporating data from multiple sources
 - Improve interoperability and data integrations
- Using data to make healthcare proactive, rather than reactive
 - Expand technology assessment to incorporate personalized healthcare
- Improving security and privacy
 - Find more efficient ways to ensure patient privacy without limiting the quality of data
- Making data from technology available
 - Improve interoperability to collect and analyze data from advancing technologies
- Improving data standardization and quality
 - Convert unstructured data into forms that are easily analyzed

Using emerging Healthcare Informatics technologies to analyze and interpret population health data is imperative in preventive medicine. We cannot strengthen preventive care until we can understand what is causing it to decline. To do this we need access to improved data.

If the data from the Healthy People 2030 project could be broken down into county, or even state data, we could collect additional data that is relevant to the population in our area, making it easier to effect changes locally. Population health problems cannot be corrected on a national or global level. Every community has different needs and challenges and until we can identify them, we cannot improve them.

References

1. Borksy, A., et al. (2018). Few Americans Receive All High-Priority, Appropriate Clinical Preventive Services. *Health Affairs*, 37(6). DOI: 10.1377/hlthaff.2017.1248
2. *Healthy People 2030 Framework*. Healthy People 2030 Framework - Healthy People 2030. (n.d.). <https://health.gov/healthypeople/about/healthy-people-2030-framework>
3. *Preventive care*. Preventive Care - Healthy People 2030. (n.d.). <https://health.gov/healthypeople/objectives-and-data/browse-objectives/preventive-care>
4. Razzak, M. I., Imran, M., & Xu, G. (2019b). Big Data Analytics for Preventive Medicine. *Neural Computing and Applications*, 32(9), 4417–4451. <https://doi.org/10.1007/s00521-019-04095-y>
5. Jothi, N., Rashid, N. A., & Husain, W. (2015). Data mining in Healthcare – A Review. *Procedia Computer Science*, 72, 306–313. <https://doi.org/10.1016/j.procs.2015.12.145>
6. Batko, K., & Ślęzak, A. (2022). The use of Big Data Analytics in Healthcare. *Journal of Big Data*, 9(1). <https://doi.org/10.1186/s40537-021-00553-4>
7. Habehh, H., & Gohel, S. (2021). Machine learning in Healthcare. *Current Genomics*, 22(4), 291–300. <https://doi.org/10.2174/1389202922666210705124359>
8. U.S. Department of Health and Human Services. (n.d.). *Data sources and methods*. Data Sources and Methods - Healthy People 2030. <https://health.gov/healthypeople/objectives-and-data/data-sources-and-methods>