

# Big Data Analysis of Home Healthcare Services

Shilpa Balan<sup>1</sup>, Joseph Otto<sup>2</sup>  
<sup>1,2</sup> *California State University – Los Angeles*

**Abstract** – The healthcare industry is rapidly changing and currently there is a constant need for improvisation because of the lack of processed data. The healthcare industry in the United States has generated vast amounts of data, but these data have not been processed effectively to make improvements in the healthcare industry. In particular, there has been an increasing demand for Home Healthcare services in the United States. This study demonstrates an application of big data methodology to study home healthcare services. The analysis is performed using statistical programming, called ‘R’, that is used for analysing big data.

**Keywords** – Big data, data, health, home healthcare, R.

## I. INTRODUCTION

The healthcare industry has generated large amounts of data [1]. Data from the U.S. healthcare system alone reached 150 exabytes in 2011 [2]; an exabyte is 1 quintillion or 10 to the 18<sup>th</sup> power bytes of data. This large amount of data is hard to process quickly and efficiently.

There is a constant need to improve the quality of healthcare. Several factors impact the need to improve healthcare quality. Amongst these are the rising rates of chronic diseases, an increase in the aging population, and a decreasing mortality rate. Driven by the potential to improve the quality of healthcare, garnering intelligence from the massive quantities of data, known as ‘big data’, can provide many insights into improving healthcare [3]–[6].

The use of home healthcare services in the United States has increased over the past 25 years [7]. The goal of home healthcare services is to help individuals live with greater independence and help the patient remain at home, therefore avoiding hospitalisation or re-hospitalisation [8]. Home healthcare services are expected to continue to grow over the next 10 years [6]. The most significant factor identified that affects home healthcare is ‘aging’ in the post-World War II era. It is predicted that by 2030, the number of people in the United States aged 65 and above will increase to 72 million [9]. This increase along with the increased risk of medication errors with the elderly could increase the number of older home healthcare patients [10].

Home healthcare is the fastest growing sector in the healthcare industry [6]. The patient population served by home healthcare is huge. Certain significant factors of the home healthcare environment impact patient safety and quality [11]. The Centre for Medicare and Medicaid Services estimates that 8 090 home healthcare agencies in the United States deliver care for more than 2.4 million elderly and disabled people annually [12]. In 1960, less than 1 million Americans were 85 years or older. By 2000, this number increased to 4.2 million,

and it is estimated that by 2030 nearly 10 million Americans will be 85 years or older [13].

With the increasing need of home healthcare required for aging citizens that now make up a huge percentage, it is important to identify aspects that affect patient health and safety [2]. When big data is analysed, healthcare providers can develop more insightful diagnoses and treatments [14], [15]. This would lead to higher quality care at lower costs and lead to better outcomes. Various big data analytic tools can be used to analyse large sets of data. In this paper, we use open source R-programming to analyse the data to determine the quality of home healthcare services in the United States.

In the following sections, we detail the background on home healthcare, research methodology, results and analysis and the discussion of our results.

## II. BACKGROUND

In this section, we discuss the background of home healthcare services in the United States. There is still limited research being conducted on home healthcare.

Home care services are assigned by a home health agency. Home health services represent a set of skilled services provided to patients at home. The home care services can be offered through nurses, therapists, social workers, and volunteers. Home care services are usually available throughout the day and at any time of the week. However, most home services are completed during the day. Through a home care agency, one can get similar services as that in a hospital [16].

Home care services can be paid for by Medicare, Medicaid, the Veterans Administration or private third-party payers such as health insurance companies or by the patient and/or their family [16]. Medicare is the largest provider for home healthcare costs [17].

There are three types of home care agencies [6]:

- Certified home health agencies (CHHAs): CHHA’s serve both Medicare and Medicaid recipients;
- Long-term home healthcare programs (LTHHCPs): Due to services provided through LTHHCP individuals eligible for nursing homes can remain at home;
- Licensed home care service agencies (LHCSAs): LHCSAs provide services such as nursing care, home health aides, personal care, private duty nursing, homemakers, and physical/occupational and speech therapies.

A large percentage of home care patients have heart disease diagnoses, followed by injuries, osteoarthritis, and respiratory ailments, requiring a greater intensity of care [18]. All these trends suggest that home care is becoming more challenging for all involved. By increasing the awareness of the health hazards

inherent in the home care environment, it may be possible to reduce the risk of injury and illness to the patient.

Home health is delivered in the United States by providers, both for-profit and non-profit. It has been estimated that there are currently more than 13 000 home health agencies serving patients across America. Approximately 9 800 of these agencies are certified to treat Medicare patients [17].

Medical errors are now thought to be a leading cause of death and further injury [19]. The risk of medication errors may be high in older home healthcare patients [20]. The promise of home health improving safety and quality and reducing medication errors has led to the hope of increasing the number of patients being able to stay in the home setting.

### III. METHODOLOGY

The home healthcare service dataset we studied is an open dataset published by the United States Department of Health & Human Services [21]. The Project Open Data enables agencies to adopt the Open Data Policy and reveal its potential [22].

The dataset describes the quality of services provided by the home health team to their patients. For example, some of the services for which the quality is indicated in the dataset are how often the home health team taught patients about their drugs, how often the home health team made sure the patients received a flu shot, and how often home health team patients had to be admitted to the hospital.

The dataset studied was converted to a suitable format in Excel. The data were cleaned to remove any missing and duplicate entries. We then used R programming to analyse the data. R is an open source program for statistical computing that is used for big data analysis [1].

This study identifies patterns for various indicators of home healthcare quality such as how often the home health team began their patient care in a timely manner and how often patients were taught about drugs. To find the median value in R of how often the home health team taught patients about their drugs, we used the median function in R as shown in Fig. 1.

```
median(homedata$howoftenhomehealthteamtaught
patientsabouttheirdrugs, na.rm=FALSE)
```

Fig. 1. Example of median function in R.

Na.rm is used to remove missing values. The output in R for the median function is:

```
[1] 0.93.
```

The Results section discusses the details of our findings from the data. Table I describes the statistical findings for all the questions of analysis.

### IV. RESULTS

Table I provides statistical findings of some of the questions we analysed from the given data. From Table I, we see that home health teams generally began treatment in a timely manner and provided appropriate instructions to the patients regarding their medicines. Home health teams also checked patients for various vaccinations and pain. Not many home healthcare patients required admission to hospitals or unplanned visits to an ER.

We were further able to analyse the data by plotting useful graphs in R. For example, one of the questions that we analysed using R: 'How often the home health team taught patients or their family caregivers about their drugs'. This question is analysed by a visualisation in R as shown in Fig. 2. Figure 2 analyses the relationship between 'how often the home health team taught patients or their family caregivers about their drugs' and 'how often patients got better'. As illustrated in Fig. 2, the quality of service provided by the home health team in this regard is very high. Further, Fig. 3 indicates that patients gave a higher quality rating when the home health team began patients' care in a timely manner.

TABLE I

STATISTICAL ANALYSIS OF HOME HEALTH SERVICE DATA

| Question of Analysis   | Median (in %) |
|--|---------------|
| How often the home health team began their patients' care in a timely manner   | 92 %          |
| How often the home health team taught patients or their family caregivers about their drugs                                | 93 %          |
| How often the home health team checked patients' risk of falling   | 99 %          |
| How often the home health team made sure that their patients had received a flu shot for the current flu season            | 73 %          |
| How often the home health team made sure that their patients had received a pneumococcal vaccine pneumonia shot            | 74 %          |
| How often the home health team checked patients for pain   | 99 %          |
| How often the home health team treated their patients' pain  | 99 %          |
| How often patients' wound improved or healed after an operation  | 90 %          |
| How often patients got better at taking their drugs correctly by mouth   | 54 %          |
| How often patients receiving home healthcare needed urgent unplanned care in the ER without being admitted to the hospital | 13 %          |
| How often home health patients had to be admitted to the hospital  | 16 %          |

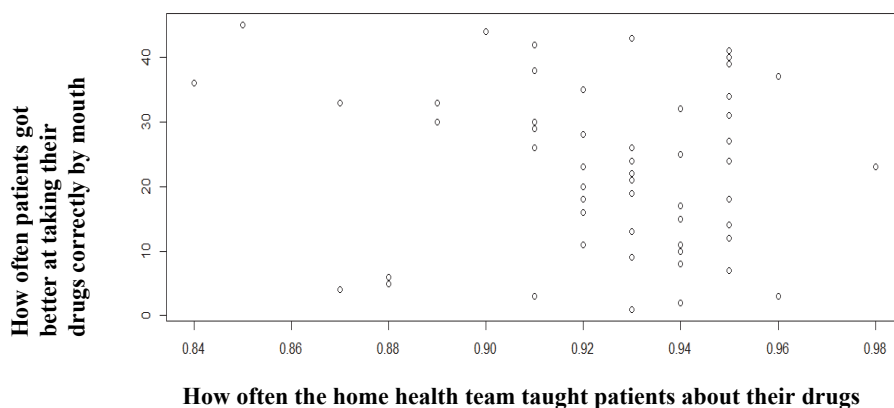


Fig. 2. Analysis of relationship between 'how often the home health team taught patients or their family caregivers about their drugs' and 'how often patients got better'.

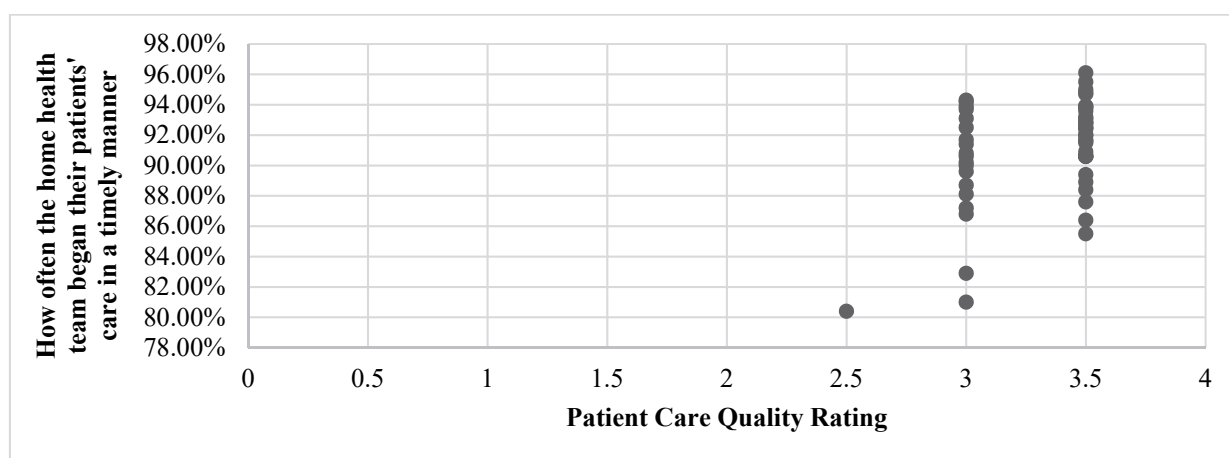


Fig. 3. Relationship between 'How often the home health team began their patients' care in a timely manner' and 'Patient Care Quality Rating'.

## V. DISCUSSION

By using R programming, one can visualise the data to better understand the big picture of a large dataset. Visuals prove to be more useful for making decisions. Static visualisations are used to support storytelling, which is the ability to tell a meaningful story using the data, usually in the form of diagrams and charts [8].

From Fig. 2, it is indicated that as the home health team taught patients about the drugs, the patients got better at taking the drugs correctly, causing less problems. Figure 3 further illustrates that as the home health team begin patients' care in a timely manner, the patient care quality rating is high.

Table I shows that a median of 13 % of the patients receiving home healthcare needed urgent, unplanned care in the ER without being admitted to the hospital. A median of 16 % of the home health patients had to be admitted to the hospital. The percentage of home health patients who need to be admitted to the hospital and ER is quite low. This indicates that the quality of services provided by the home health team is outstanding and effective.

## VI. CONCLUSION

Each and every healthcare sector has its own problems, opportunities and challenges. Healthcare data are expected to continue to grow further in the years ahead [23]. Our data analysis points out that the home health team is delivering high quality care to patients in their homes.

Visualisation research has traditionally focused on the exploration and analysis of data [24]. As a result of the growth in the population of older Americans, home healthcare is growing in relevance. Moreover, many of the older Americans strive to remain independent. These demographic trends show the increasing importance of home healthcare. Home healthcare programs are also now engaging in emergency response planning [25].

Home healthcare is now successfully treating more critical health conditions. Further research can be conducted on the challenges faced by home healthcare programs for different health conditions of their patients.

## REFERENCES

- [1] R Project. The R Project for Statistical Computing. [Online]. Available: <https://www.r-project.org/> Accessed: Feb. 15, 2016.
- [2] R. Gershon, M. Pogorzelska, K. Qureshi, P. Stone, S. Samar, L. Westra, M. Damsky and M. Sherman, "Home Health Care Patients and Safety Hazards in the Home: Preliminary Findings," *Advances in Patient Safety: New Directions and Alternative Approaches*, vol. 1, pp. 1–16, 2008.
- [3] C. Burghard, "Big Data and Analytics Key to Accountable Care Success," *IDC Health Insights*, 2012.
- [4] A. Dembosky, "Data Prescription for Better Healthcare," *Financial Times*, 2012. [Online]. Available: <http://www.ft.com/intl/cms/s/2/55cbca5a-4333-11e2-aa8f-00144feabdc0.html#axzz2W9cuwajK> Accessed: Feb. 5, 2016.
- [5] B. Feldman, E. M. Martin and T. Skotnes, Big Data in Healthcare Hype and Hope, 2012. [Online]. Available: <http://www.west-info.eu/files/big-data-inhealthcare.pdf> Accessed: Feb. 18, 2016
- [6] L. Fernandes M. O'Connor and V. Weaver, *Big data, bigger outcomes*, 2012.
- [7] B. C. Vladeck and N. A. Miller, "The Medicare home health initiative," *Health Care Finance Rev.* vol. 16, issue 1, pp. 7–16, 1994.
- [8] E. Segel and J. Heer, "Narrative Visualization: Telling Stories with Data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 6, pp. 1139–1148, Nov.-Dec. 2010. <https://doi.org/10.1109/TVCG.2010.179>
- [9] C. H. Ellenbecker, L. Samia, M. J. Cushman and K. Alster, "Patient Safety and Quality in Home Health Care. Patient Safety and Quality," in *Patient Safety and Quality: An Evidence-Based Handbook for Nurses*, R. G. Hughes, Ed. Rockville, MD: Agency for Healthcare Research and Quality, 2008, ch 13.
- [10] Medicare and Medicaid statistical supplement. *Health Care Finance Rev.* Statistical Supplement. pp. 126–141, 1997.
- [11] C. J. DeFrances, M. J. Hall and M. N. Podgornik, 2003 National hospital discharge survey. [Online]. Available: [www.cdc.gov/nchs/data/ad/ad359.pdf](http://www.cdc.gov/nchs/data/ad/ad359.pdf) Accessed: Apr. 12, 2008.
- [12] Centers for Medicare & Medicaid Services. What is home health care?, 2006. [Online]. Available: <http://www.medicare.gov/HHCCompare/Home.asp?dest=NAV|Home|About|WhatIs#TabTop> Accessed: March 10, 2016.
- [13] Federal interagency forum on aging-related statistics. *Older Americans update: Key indicators of well-being*. Washington, DC: U.S. Government Printing Office, 2006.
- [14] Knowledgegent. Big Data and Healthcare Payers, 2013. [Online]. Available: <http://knowledgegent.com/mediapage/insights/whitepaper/482> Accessed: March 12, 2016.
- [15] W. Raghupathi, "Data Mining in Health Care," in *Healthcare Informatics: Improving Efficiency and Productivity*. Boca Raton, FL: CRC Press., 2010, ch 11. <https://doi.org/10.1201/9781439809792-c11>
- [16] American Cancer Society. Home Care Agencies, 2016. [Online]. Available: <http://www.cancer.org/treatment/findingandpayingfortreatment/choosingyourtreatmentteam/homecareagencies/home-care-agencies-toc> Accessed: Apr. 10, 2016.
- [17] AHHQI. "Home Health Care: An Essential Solution to America's Health Care Challenges. The Home Health Solution," *Alliance for Home Health Quality and Innovation*, 2016.
- [18] National Association for Home Care & Hospice. Basic statistics about home care, 2015. [Online]. Available: [www.homecareaware.org](http://www.homecareaware.org) Accessed: Nov. 1, 2015.
- [19] Committee on Quality of Health Care in America. *To Err Is Human: Building a Safer Health System*. Institute of Medicine. Washington, D.C: National Academy Press, 2000. <https://doi.org/10.17226/9728>
- [20] S. Meredith, P. Feldman and D. Frey, "Improving medication use in newly admitted home health care patients: a randomized controlled trial," *Journal of the American Geriatrics Society*, vol. 50, issue 9, pp. 1484–1491, Sep. 2002. <https://doi.org/10.1046/j.1532-5415.2002.50402.x>
- [21] Home Health Care National Data. Home Health Care National Data. Department of Health & Human Services. [Online]. Available: <http://www.healthdata.gov/dataset/home-health-care-hhcahps-national-data> Accessed: Jan., 2016.
- [22] Project Open Data. Open Data Policy- Managing Information as an Asset, 2016. [Online]. Available: <https://project-open-data.cio.gov/>
- [23] IHTT. Transforming Health Care through Big Data Strategies for leveraging big data in the health care industry, 2013. [Online]. Available: <http://ihealthtran.com/wordpress/2013/03/ih%C2%B2-releases-big-data-research-reportdownload-today/> Accessed: Feb. 18, 2016.
- [24] R. Kosara and J. Mackinlay, Storytelling: The Next Step for Visualization. 2016. [Online]. Available: [http://kosara.net/papers/2013/Kosara\\_Computer\\_2013.pdf](http://kosara.net/papers/2013/Kosara_Computer_2013.pdf)
- [25] A. Knebel and S. J. Phillips, Eds., *Home Health Care during an Influenza Pandemic: Issues and Resources*. AHRQ Publication No. 08-0018. Rockville, MD: Agency for Healthcare Research and Quality, 2008.

**Shilpa Balan** is an Assistant Professor of the Department of Information Systems at California State University – Los Angeles. Her research interests include big data, business intelligence and healthcare informatics. E-mail: [sbalan@calstatela.edu](mailto:sbalan@calstatela.edu)

**Joseph Otto** is the Chair and Professor of the Department of Information Systems at California State University – Los Angeles. His research interests include big data analytics, business intelligence and visualisation. E-mail: [jotto@calstatela.edu](mailto:jotto@calstatela.edu)