

Prevalence of sociodemographic factors in a cohort of diabetes mellitus: a retrospective study

Madurapperumage, A, Erandathi
University of Waikato

Michael Mayo
University of Waikato, Hamilton, New Zealand

William Y.C.Wang*
University of Waikato, Hamilton, New Zealand

Ibrahim Shafiu
Waikato District Health Board, New Zealand

ABSTRACT

Exploring the sociodemographic factors of a cohort is a vital phase in revealing significant aspects of the societal health status. The health care sector utilises the results of exploratory analysis of the sociodemographic nature to fulfil various purposes such as constructing health care policies, allocating adequate resources, imposing necessary medications and many more. A large and growing body of evidence shows that understanding the pervasiveness of sociodemographic factors: age, ethnicity, gender, reveal crucial information. Therefore, this study aims to disclose the knowledge through analysing the sociodemographic details of a New Zealand diabetes cohort. Diabetes mellitus is a chronic fatal disease that occurs due to the inability to control proper blood sugar levels, which causes multitudinous acute and chronic complications. Diabetes became a high prevalence disease in the region of Waikato. Analysing the cohort of diabetes patients associated with complications of diabetes illustrate the prevalence of complications of diabetes among the patients. The dataset of the study has been collected from the Waikato district health board. This study intends to report the initial scanning of the dataset profile with visualising the resulting patterns of sociodemographic details from the samples and their association with complications of diabetes. The Sankey diagrams use to visualise the results of exploratory data analysis. The resulted graphs of the data screening descriptively illustrate the characteristics of the cohort associated with demographic factors. Maori population shows higher percentage (0.68) of diabetes patients than the other ethnicities, while having narrower age expansion (13-95) with early onset age, compared to others (20-103). Males (0.61%) are more vulnerable to diabetes than females (0.55%). Additionally, hypertension and cardiovascular diseases are common among the diabetes patients' in the Waikato region. Maori male population is highly vulnerable to diabetes. This study will be beneficial in constructing and analysing the demographical categories of the cohort to comparatively study the pervasiveness of the diseases among resulting classes.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ICMHI 2022, May 13–15, 2022, Virtual Event, Japan

© 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-9630-1/22/05...\$15.00
<https://doi.org/10.1145/3545729.3545768>

CCS CONCEPTS

• Applied computing; • Life and medical sciences; • Health care information systems;

KEYWORDS

Diabetes mellitus, Complications of diabetes, Exploratory data analysis, Sociodemographic features

ACM Reference Format:

Madurapperumage, A, Erandathi, William Y.C.Wang*, Michael Mayo, and Ibrahim Shafiu. 2022. Prevalence of sociodemographic factors in a cohort of diabetes mellitus: a retrospective study. In *2022 6th International Conference on Medical and Health Informatics (ICMHI 2022)*, May 13–15, 2022, Virtual Event, Japan. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3545729.3545768>

1 INTRODUCTION

Diabetes mellitus is one of the fastest growing global pandemic in the 21st century. There were 537 million diabetes patients around the globe in 2021, which has been projected to reach 643 million in 2030 and 783 million by 2045. Further, the mortality of diabetes or its complication has been estimated at 6.7 million adults in 2021 (IDF, 2021). Diabetes is a chronic metabolic disorder which caused by insulin deficiency in the body due to the dysfunctions of beta cells, insulin resistance or both (Cole & Florez, 2020). Although there are three main types of diabetes: type 1 (T1DM), type 2 (T2DM), and gestational diabetes (GDM), the most common type of diabetes is T2DM which accounts for 90% of diagnoses of diabetes globally. In T2DM, body cells are resistant to insulin, leading to excessive production of the hormone. Eventually, pancreatic beta cells, which produce the hormone, fail to meet demand (Forouhi & Wareham, 2010). As a result, the blood glucose level is high in patients with T2DM.

In considering the global picture of the distribution of diabetes mellitus, a 27% increase is estimated in the Western Pacific region in 2021- 2045 (IDF, 2021). New Zealand has been experiencing a statistically significant increase in diabetes patients over the last decade, estimated at 3.4% to 4.0%, where 277,800 people in Aotearoa, New Zealand were estimated as diabetes patients in 2021 (NZ Ministry of Health, 2022). Additionally, the estimated diabetes-related expenditure of New Zealand per person in 2021 is 3967.5 USD (IDF, 2021). The high prevalence rate, the enormous impact of diabetes on the country's health status, and the high expenditures lead to exploring the information of diabetes cohort of New Zealand. As a result, the disclosed information may provide insights into policymaking and health care management.

Exploring the sociodemographic features of a country's cohort is a vital step in understanding the behaviours and existing patterns, which can be directly utilised in the advancement of the management of the health care sector. Scholars use a spectrum of sociodemographic features to reveal the information of cohorts. For example, level of education, civil status, age and sex was used as sociodemographic factors in a study finding the association with HbA1c variability among T2DM patients (Mellergård, Johnsson, & Eek, 2020). (Willers, et al., 2018) used sex, age, educational level, marital status, and region of birth as sociodemographic features when determining the sociodemographic determinants and health outcome variation in a cohort with T1DM. In a study on seeking health-related quality of life in diabetes associated with its social and clinical determinants, age, sex, occupation, education level, and marital status are considered as sociodemographic features to measure the quality of life of diabetes patients (Rodríguez-Almagro, García-Manzanares, Lucendo, & Hernández-Martínez, 2018). According to (Young Hyman, et al., 2016), sex and age are two crucial factors influencing the prevalence of diabetes either independently or as covariates. Further, ethnicity has been widely used as a feature to report and categorise the diabetes cohorts (IDF, 2021), (WHO, 2016). Usage of the characteristics of different cohorts in different studies clearly expresses the importance of ethnicity-specific benchmarks. The studies conducted in multi-ethnic countries usually consider ethnicity as a crucial sociodemographic factor for differentiating and specifying the characteristics of different ethnic groups (Kyrrou, et al., 2020), (Boakye, et al., 2018). Further, (Mellergård, Johnsson, & Eek, 2020), point that sociodemographic factors associated with diabetes: age, ethnicity, gender vary significantly between countries. The exploratory analysis of sociodemographic factors of a diabetes cohort provides insights for the health care management sector.

Furthermore, diabetes leads to a range of fatal macro and microvascular complications. The most common complications of diabetes are heart disease, stroke and hypertension, blindness, kidney disease, nervous system disorders, amputation, pregnancy problems, other complications (Scientific American, 2022). According to (WHO, 2016), diabetic retinopathy caused 2.6% of blindness in 2010, whereas 80% of end-stage renal diseases are consequences due to diabetes. Further, adults with diabetes have a two or three-times higher rate of cardiovascular risk than those without diabetes. Furthermore, the lower limb amputation incidents are drastically increased among the population with diabetes over the past decade. The data shows that it reached 3.5 events per 1000 persons diagnosed with diabetes per year in 2016 (WHO, 2016). A study of the epidemiology of diabetes and its complications revealed that up to 65% of diabetes deaths are due to cardiovascular diseases, 44% of new cases of end-stage renal diseases are due to diabetes, and 30%-50% of diabetes patients are suffering from peripheral neuropathy (Deshpande, Harris-Hayes, & Schootman, 2008). Therefore, the distribution of complications in a diabetes cohort is another way of understanding the underneath image of a cohort.

This study focuses on exploring the sociodemographic factors of a diabetes cohort in New Zealand while understanding the distribution of complications of diabetes. The need of more studies on the sociodemographic features among diabetes patients in the Waikato region, less exploration of the distribution of complications

of diabetes, and the importance that the result of an exploratory analysis for the health care sector made the desire of this study.

2 METHODOLOGY:

This study was undertaken using a data set collected from the Waikato district health board (WDHB). Random sampling has been used in collecting the dataset. The national health indexes are converted to patient reference numbers to protect the privacy of the patients. The pre-processed dataset consists of 2656 diabetes patients from 2018 to 2020 in the Waikato region. Population details of the Waikato region are collected from the official website of Stats NZ (Waikato Region, 2022). The exploratory analysis of the dataset was conducted using the Python programming language. The sociodemographic factors used in this study are age, sex, and ethnicity. Due to the retrospective design study and the existing limitations of data issuing policies of WDHB, the sociodemographic features have to be restricted. Data pre-processing techniques have been used in data cleaning and data transformation. The patients with missing attributes such as age and the noisy values were removed from the dataset to maintain consistency. In contrast, values with a negligible amount of data were removed to avoid the unnecessary complexity of the results. The data frames are created to fulfil the requirement of exploratory analysis by extracting and aggregating the necessary attributes. Although the sociodemographic factors have to be limited due to the retrospective nature of the study, the selected features are matched with existing scholarly works. Diabetes complications have been extracted from the dataset to understand the distribution of micro and macrovascular complications of diabetes among the cohort. Among the unique disease set of the dataset, the diseases with the highest number of patients has been chosen to extract the complications of diabetes among the cohort. The results of exploratory data analysis have been presented using tables and a Sankey diagram, to visualise the results. The complications with more than 80 patients are selected to visualise in the Sankey diagram to simplify the representation.

3 RESULTS:

The dataset collected from WDHB consists of details of diabetes patients from 2018 to 2020 who lives in the Waikato region. There are 2656 diabetes patients who have been selected for the study after pre-processing the dataset. The prevalence of diabetes in the Waikato region is 5.79%. The characteristics of all patients are tabulated in Table1. The selected cohort consist of more males (52%) than females (48%). Moreover, the cohort comprises twenty-two ethnic groups, categorised into five classes in this study: European, Maori, Pacific, Asian, and others. The majority of the patients in this cohort are Europeans, followed by Maori.

The age distribution of each group of patients is represented in the Table2. The mean age of diagnosis of diabetes is always less for females. But the age dispersion of females is wider in overall and Non-Maori groups. Mean age of diagnosis and age expansion are less among Maori than Non-Maori. The earliest diagnosis mean of age belongs to the Pacific ethnic group, whereas Europeans show the oldest mean diagnosed age. Maori possesses the minimum diagnosis age (13), and the maximum (103) belongs to Europeans.

Table 1: Characteristics of diabetes patients in the Waikato region (N=2656)

Factors	Categories	N=2656	Percentage	Patient/ Population percentage
Gender	Male	1380	51.96	0.61
	Female	1275	48.00	0.55
Ethnicity	European (NZ European/Pakeha, Other European, European No Further Definition)	1649	62.09	0.48
	Maori	744	28.01	0.68
	Pacific (Cook Island Maori, Samoan, Fijian, Tongan, Other Pacific Peoples, Niuean, Pac People No Further Definition)	92	3.46	0.44
	Asian (Indian, Other Asian, Chinese, South-East Asian, Asian No Further Definition)	131	4.93	0.30
	Other (Not Stated, African, Other Ethnicity, Middle Eastern, Latin American/Hispanic, Response Unidentifiable)	40	1.51	0.38

Table 2: Age distribution of diabetes cohort in identified sociodemographic groups

	Categories	Mean (Standard deviation)	Minimum value(0th)	First quartile (25 th)	Median(50 th)	Third quartile (75 th)	Maximum value (100 th)
All patient		67.4(14.2)	13.0	58.0	69.0	78.0	103.0
Gender	Male	69.3(13.1)	23.0	61.0	71.0	79.0	97.0
	Female	66.6(15.1)	13.0	57.0	68.0	78.0	103.0
Maori	Male	61.5(13.1)	23.0	53.0	62.0	70.0	95.0
	Female	60.3(14.5)	13.0	52.0	62.0	70.0	91.0
Non-Maori	Male	71.8(12.1)	33.0	64.0	73.0	80.0	97.0
	Female	69.7(14.5)	20.0	61.0	72.0	80.0	103.0
Ethnicity	European	71.7(12.5)	20.0	64.0	73.0	80.0	103.0
	Maori	60.2(13.9)	13.0	52.0	61.0	69.0	93.0
	Pacific	60.0(15.0)	26.0	50.0	59.0	71.0	92.0
	Asian	61.0(15.0)	29.0	49.0	63.0	72.0	94.0
	Other	66.7(14.2)	24.0	60.3	68.0	75.0	92.0

The dispensation of microvascular and macrovascular complications of the cohort of diabetes mellitus is represented in figure 1, with their ethnicities. The left side of the diagram shows the ethnicity expansion, while the right-hand side represents the distribution of frequent complications of the cohort. The most common complications are shown in the Sankey diagram which has at least 99 patients (3.73%) of the cohort. The most frequent complication among the patient is hypertension, followed by T2DM with features of insulin resistance, kidney complication, ophthalmic complication, and atherosclerotic heart disease. Almost half of the cohort have hypertension, where 23.5%, 18.9%, 13.1% have kidney complications, ophthalmic complications, and neuropathy, respectively. The results of this analysis show a consonant with the overall distribution of complications over the globe (WHO, 2016).

4 CONCLUSION:

Exploring the demographic factors of a cohort is vital in revealing crucial information about a nation. Due to the strong influence of diabetes mellitus in the health care sector of New Zealand, analysing the demographic profile reveals various perspectives of the nation. The results of this study show that the Maori ethnic group has a higher patient population percentage than other ethnic groups (0.68%) in the Waikato region. Although more males (0.61%) are in diabetes than females (0.55%), the average ages of diagnosis are less in females than males. Further, the diagnosis age of diabetes among Maori (60.2) is less than Non-Maori (70.3). The earliest age of diagnosis can be seen among Maori (60.2) and Pacific (60.0) people, where the highest average age belongs to Europeans (71.7). The results of this research coincide with the overall statistics of New Zealand (NZ Ministry of Health, 2022). The most common

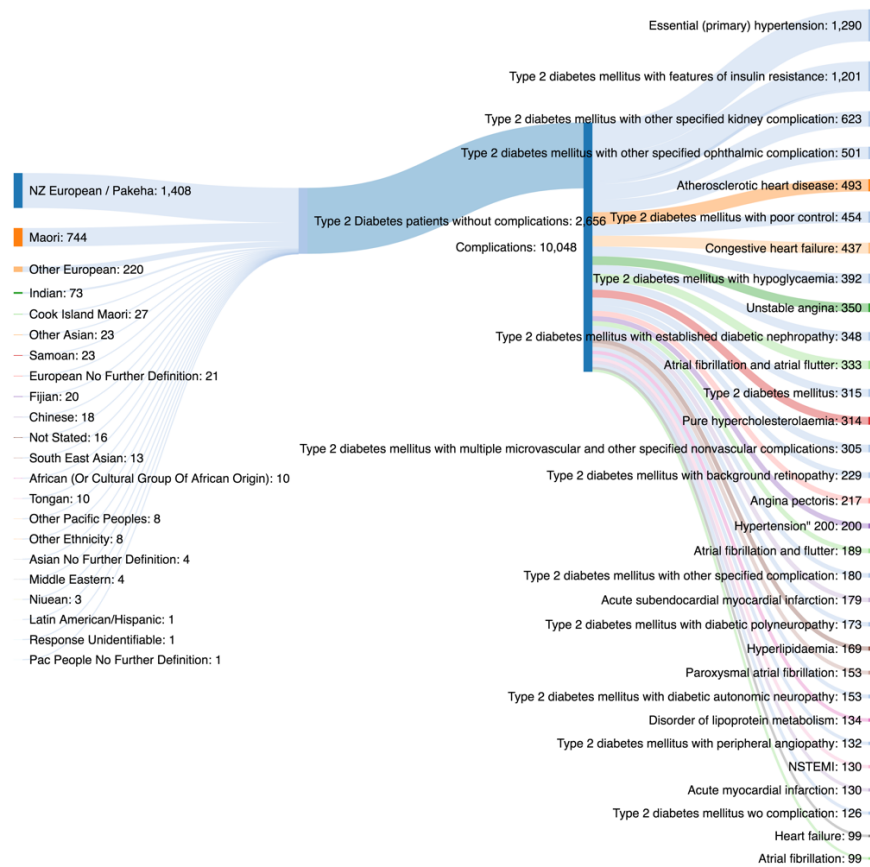


Figure 1: Distribution of complications of diabetes mellitus and the ethnicities of the cohort.

complication among diabetes patients is hypertension (n=1290, 48.6%).

Our study shows the epidemiology of a group of patients in the Waikato region, which results in exploratory analysis and an expansion of complications of diabetes mellitus. There is a scarcity in researching for distribution of diabetes complications in a cohort. The complications of diabetes have been researched under various thematic areas: epidemiology and pathogenesis, micro, macro, and miscellaneous vascular complications, treatment options (Papathodorou, 2017). Although the epidemiology, the expansion of sociodemographic factors and the association of epidemiology with specific clinical attributes such as HbA1c are commonly researched, few scholarly works can be found in researching of distribution of the complications of diabetes mellitus.

However, we note some limitations of our study. First, due to the limited number of available features in the dataset collected from WDHB, the sociodemographic characteristics analysed in the study have to be restricted. In addition, the considered time frame comprises only three years. Therefore, the more expanded historical data of the patients may reveal different cohort perspectives. Moreover, the study's findings may slightly differ from the overall statistics of the nation due to the characteristics of the considered specific cohort.

In conclusion, the Maori population in the Waikato region are more vulnerable to diabetes than Non-Maori. Additionally, females are diagnosed with diabetes earlier than males. The findings of this study can be utilised in the health care sector. The results of exploratory data analysis provide clear guidance in policymaking, resource allocating, and clustering diabetes patients. The stakeholders in the health care sector can be more attentive to highly vulnerable classes of patients. At the same time, the awareness programs and focus groups of research studies can be focused on the identified risk groups. Further, this study shows a cross-sectional diabetes cohort of the Waikato region, which consists of high diabetes densified population. A time-series analysis of the prevalence of diabetes with the evolving sociodemographic factors for a considerably long time (10 years) will reveal how the disease and its complications have evolved and the effect of changes in their sociodemographic factors in different patient groups. An extension of this study will focus on the aims mentioned above.

REFERENCES

[1] Boakye, E. A., Varble, A., Rojek, R., Peavler, O., Trainer, A. K., Osazuwa-Peters, N., & Hinyard, L. (2018). Sociodemographic Factors Associated With Engagement in Diabetes Self-management Education Among People With Diabetes in the United States. *Public Health Reports*, 685-691.

- [2] Cole, J. B., & Florez, J. C. (2020). Genetics of diabetes mellitus and diabetes complications. *Nature Reviews | Nephrology*, 377-390.
- [3] Deshpande, A. D., Harris-Hayes, M., & Schootman, M. (2008). Epidemiology of Diabetes and Diabetes-Related Complications. *Physical Therapy*, 1254-1264.
- [4] Forouhi, N. G., & Wareham, N. J. (2010). Epidemiology of diabetes. *Medicine*, 602-606.
- [5] IDF. (2021). *IDF Diabetes Atlas*. Brussels: International Diabetes Federation.
- [6] Kyrou, I., Tsigos, C., Mavrogianni, C., Cardon, G., Stappen, V. V., Latomme, J., . . . S. (2020). Sociodemographic and lifestyle-related risk factors for identifying vulnerable groups for type 2 diabetes: a narrative review with emphasis on data from Europe. *BMC Endocrine Disorders*.
- [7] Møllergård, E., Johnsson, P., & Eek, F. (2020). Sociodemographic factors associated with HbA1c variability in type 2 diabetes: a prospective exploratory cohort study. *BMC Endocrine Disorders*.
- [8] NZ Ministry of Health. (2022, February 9). *Living well with diabetes*. Retrieved from Ministry of Health NZ: <https://www.health.govt.nz/nz-health-statistics/health-statistics-and-data-sets/diabetes-data-and-stats>
- [9] Papatheodorou, K. B. (2017). Complications of Diabetes 2017. *Journal of diabetes research*.
- [10] Rodríguez-Almagro, J., García-Manzanares, Á., Lucendo, A. J., & Hernández-Martínez, A. (2018). Health-related quality of life in diabetes mellitus and its social, demographic and clinical determinants: A nationwide cross-sectional survey. *Journal of Clinical Nursing*, 4212-4223.
- [11] Scientific American, A. D. (2022, February 10). *Complications: The Toll of Diabetes*. Retrieved from SCIENTIFIC AMERICAN: <https://www.scientificamerican.com/article/complications-the-toll-of-diabetes/>
- [12] *Waikato Region*. (2022, Feb 12). Retrieved from Stats NZ: <https://www.stats.govt.nz/tools/2018-census-place-summaries/waikato-region>
- [13] WHO. (2016). *GLOBAL REPORT ON DIABETES*. France: World health organisation.
- [14] Willers, C., Iderberg, H., Axelsen, M., Dahlström, T., Julin, B., Leksell, J., . . . Lilja, M. (2018). Sociodemographic determinants and health outcome variation in individuals with type 1 diabetes mellitus: A register-based study. *PLOS One*.
- [15] Young-Hyman, D., Groot, M. d., Hill-Briggs, F., Gonzalez, J. S., Hood, K., & Peyrot, M. (2016). Psychosocial Care for People With Diabetes: A Position Statement of the American Diabetes Association. *Diabetes Care*, 2126-2140