Using electronic health record data for chronic disease surveillance in low- and middle-income countries: the example of hypertension in rural Guatemala

Sean Duffy,[1] Juan Aguirre Villalobos,[1] Alejandro Chavez,[2] Kaitlin Tetreault,[3] Do Dang,[4] Guanhua Chen,[3] Taryn McGinn Valley[1,5]

[1]Department of Family Medicine and Community Health, University of Wisconsin School of Medicine and Public Health, Madison; [2]University of California, San Francisco School of Medicine; [3]Department of Biostatistics and Medical Informatics, University of Wisconsin School of Medicine and Public Health, Madison; [4]University of Maryland Medical Center Family Medicine Residency, College Park; [5]Department of Anthropology, University of Wisconsin, Madison, USA

Correspondence: Sean Duffy, Department of Family Medicine and Community Health, University of Wisconsin School of Medicine and Public Health, 1002 Gilson St, Madison 53715, WI, USA.
Tel.: +1.608.354.7930.
E-mail: sean.duffy@fammed.wisc.edu

Key words: hypertension, EHR, LMIC, low-resource settings, disease surveillance.

Contributions: SD assumed the lead role in study design, data review, descriptive analysis, manuscript drafting and editing, and contributed to the literature review and statistical analysis; JA led the literature review and contributed to study design, data review, and descriptive analysis, manuscript drafting and analysis; AC was primarily responsible for data cleaning and preparation and contributed to study design and manuscript drafting and editing; KT and GC were primarily responsible for developing and carrying out the statistical analysis plan and contributed to manuscript drafting and editing; DD contributed to study design, data review and descriptive analysis, and manuscript drafting and editing; TV contributed to study design and manuscript drafting and editing. All the authors have read and approved the final version of the manuscript and agreed to be held accountable for all aspects of the work.

Conflict of interest: the authors declare no potential conflict of interest.

Funding: the research reported in this publication was supported by the Fogarty International Center of the National Institutes of Health under award number R21TW011891. 100% of the estimated $5239 cost of this project was supported by this grant. During the research period, TMV was funded in part by the NIH Medical Scientist Training Program grant T32 GM140935. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The NIH had no involvement in study design or execution, nor the decision to submit for publication.

Ethics approval and consent to participate: research was performed in accordance with the Declaration of Helsinki. This project was reviewed and approved by the San Lucas Healthcare Committee, as well as the University of Wisconsin Minimal Risk IRB (submission ID 2021-1164), who determined this project to represent exempt human subjects research involving review of medical records alone and as such, obtaining informed consent was not required.

Consent for publication: not applicable.

Availability of data and materials: the datasets generated and/or analyzed during the current study are not publicly available as such sharing outside of the research team was not part of the research protocols approved by ethical oversight committees for this project, but may be made available upon reasonable request from the corresponding author with IRB permission.

Acknowledgments: we would like to acknowledge the San Lucas Mission Healthcare Program, particularly Medical Director Dr. Rafael Tun, for providing access to the EHR data for this project.

Received: 8 February 2024.
Accepted: 19 June 2024.
Early access: 21 July 2024.

Healthcare in Low-resource Settings 2024; 12:12370
doi:10.4081/hls.2024.12370

Abstract

Hypertension is the leading preventable cause of death worldwide. Two-thirds of people with hypertension live in Low- and Middle-Income Countries (LMIC). However, epidemiological data necessary to address the growing burden of hypertension and other Non-Communicable Diseases (NCDs) in LMICs are severely lacking. Electronic Health Records (EHRs) are an emerging source of epidemiological data for LMICs, but have been underutilized for NCD monitoring. The objective of this study was to estimate the prevalence of hypertension in a rural Indigenous community in Guatemala using EHR data, describe hypertension risk factors and current treatment in this population, and demonstrate the feasibility of using EHR data for epidemiological surveillance of NCDs in LMIC. We conducted a cross-sectional analysis of 3646 adult clinic visits. We calculated hypertension prevalence using physician diagnosis, antihypertensive treatment, or Blood Pressure (BP) ≥140/90 mmHg. We noted antihypertensives prescribed and BP control (defined as BP<140/90 mmHg) for a total of 2496 unique patients (21% of whom were men). We constructed mixed-effects models to investigate the relationship between BP and hypertension risk factors. The estimated hypertension prevalence was 16.7%. Two-thirds of these patients had elevated BP, but were not diagnosed with or treated for hypertension. Most patients receiving treatment were prescribed monotherapy and only 31.0% of those with recognized hypertension had controlled BP. Male sex, older age, increasing weight, and history of hypertension were associated with increasing systolic BP, while history of hypertension, history of diabetes, and increasing weight were associated with increasing diastolic BP. Using EHR data, we estimated comparable hypertension prevalence and similar risk factor associations to prior studies conducted in Guatemala, which used traditional epidemiological methods. Hypertension was underrecognized and undertreated in our study population, and our study was more efficient than traditional methods and provided additional data on treatment and outcomes; insights gleaned from this analysis were essential in developing a sustainable intervention. Our experience demonstrates the feasibility and advantages of using EHR-derived data for NCD surveillance and program planning in LMICs.

Introduction

Though the COVID-19 pandemic starkly illustrated the significant impact infectious diseases continue to have on human health and flourishing, Non-Communicable Diseases (NCDs) now account for nearly three-quarters of all deaths globally.[1] Hypertension, the leading preventable cause of death and disability worldwide, is an exemplar NCD. An estimated 1.13 billion people around the world currently live with hypertension, with two-thirds of these patients living in Low- and Middle-Income Countries (LMIC).[2,3] Screening and treatment of hypertension have been found to be cost-effective in reducing morbidity and mortality across a broad range of settings.[4] Despite this, less than 40% of patients with hypertension in LMIC are aware of their condition and less than 10% have good control of their Blood Pressure (BP).[3] Furthermore, health systems in LMIC, which are often focused on providing episodic care for acute illnesses and suffer from inadequate and poorly distributed health care infrastructure and workforce, are ill-equipped to address the ongoing rise in NCDs, such as hypertension.[5-7]

A key barrier to combating hypertension and other NCDs in LMICs is a lack of information about current prevalence and control — particularly at the local level in rural and resource-limited areas — which is necessary to design, implement, and monitor effective health programs.[8] Despite shouldering a significant majority of the global burden of disease, funding for such research in LMIC settings makes up only a small proportion of global research funding.[9,10] At the same time, implementation of Electronic Health Record (EHR) systems is increasing in LMIC.[11] Using data collected with such systems in the course of routine care for disease surveillance and program monitoring has been proposed as an alternative to traditional, population-survey-based methods, which are resource- and time-intensive and thus not well-suited to areas with already insufficient health care funding and workforce.[12] However, this EHR-based approach has thus far mostly been applied to infectious disease and maternal health and there are few examples of its use in addressing NCDs.[12]

Guatemala is a middle-income country where the burden of NCDs is increasing, accounting for 59% of deaths in recent reports.[13] Though hypertension plays a role in these deaths, estimates of the prevalence of hypertension in Guatemala are scarce, particularly for rural Indigenous communities. In addition, little is known about the risk factors for hypertension and the proportion of those with hypertension who have been diagnosed and are being treated in these populations. For nearly a decade, our team has been working with Guatemalan physicians and Community Health Workers (CHWs) to improve capacity for primary care, chronic disease management, and prevention in San Lucas Tolimán, a rural municipality in the Western Highlands region of Guatemala. San Lucas has a population of approximately 30,000 people, of which a majority belong to the Kaqchikel Maya Indigenous group.[14] In addition to a small hospital run by our partner organization, the San Lucas Mission (SLM), and a clinic and a few small health outposts run by the government, the community receives medical care from visiting medical teams organized by SLM. These teams, mostly from the United States on one-week to one-month trips, operate mobile medical clinics in the rural communities of San Lucas on a rotating basis throughout most years. In general, access to healthcare services in rural communities is poor, and these periodic mobile clinics are often the only formal healthcare that these patients receive.

A key component of our work in Guatemala has been the digitization of existing workflows and integration of digital and Mobile Health (mHealth) elements into new programs to improve the reliability and efficiency of recordkeeping, enhance CHW capabilities, and facilitate program monitoring and effectiveness evaluation. To this end, in 2018, our team developed and implemented an EHR to replace the existing paper system. This EHR uses the Internet-in-a-Box platform on Raspberry Pi devices to create a local Wi-Fi network.[15] Clinicians use their own mobile devices or laptops to connect to this network to view and update existing medical records, as well as create records for new patients. The EHR uses multiple choice responses or validated data field (*e.g.* numeric) entry instead of free text input whenever possible, which presents several advantages. First, it supports toggling between English and Spanish to facilitate review by both visiting and Guatemalan personnel. Secondly, it allows for the provision of clinical decision support through the use of standard diagnosis codes and suggested treatments based on diagnosis, local standard of care, and formulary. Finally, these features facilitate data collection and analysis for epidemiologic and program monitoring.

In 2020, we started planning to adapt our successful mHealth-based and CHW-led diabetes program[16] to hypertension and to robustly evaluate this new intervention. Accurate data on the current burden of disease and treatment of hypertension in these communities, which was sorely lacking, was necessary to develop an effective program to identify and treat patients with hypertension. To alleviate the lack of primary research investment in these communities, we turned to the newly implemented EHR as a novel source of epidemiologic data. Accordingly, the objectives of this study were to i) estimate the prevalence of hypertension among adults living in the rural Indigenous communities of San Lucas Tolimán (SLT), Guatemala using EHR-derived data; ii) elucidate risk factors for hypertension among the population; iii) describe the medical treatment and adequacy of BP control for patients diagnosed with hypertension; and iv) demonstrate the feasibility of using EHR data for epidemiological surveillance of chronic NCDs in low-resource settings.

Materials and Methods

Data collection, inclusion and exclusion criteria

For this cross-sectional study, we used deidentified data from the SLM database of patients seen by visiting medical teams in mobile clinics from late May 2018, when the EHR was first deployed, through early March 2020, when mobile clinics were suspended due to the COVID-19 pandemic. Health records of patients 18 years and older were searched to obtain the following information for analysis: age, sex, BP, weight, height, diagnoses, and medication treatment. We excluded patients younger than 18 years of age and any records without a listed age. We reviewed data for possible erroneous entries (for example, diastolic BP greater than systolic BP) and excluded these from our analysis.

We used the STROBE checklist for cross-sectional studies,[17] available in supplementary materials.

Case definitions and calculation of hypertension prevalence

The denominator we used for the calculation of hypertension prevalence was the number of unique patients with at least one clinic visit with a listed diagnosis. The rationale for using this denominator was to increase the likelihood of counting only completed visits, as patients without a listed diagnosis may have registered for a visit, but may not have actually been seen. We used two different figures, the number of confirmed hypertension cases and the number of possible hypertension cases, and their sum as numerators for the prevalence calculations.

We defined a “confirmed hypertension case” as any patient with a listed diagnosis of hypertension or a prescription for an antihypertensive medication during at least one visit. We determined hypertension diagnosis based on both hypertension diagnoses selected from the preset EHR options as well as diagnoses consistent with hypertension entered as free text (as determined by manual review). We determined whether a patient had received an antihypertensive treatment in the same manner. We defined a "possible hypertension case" as a patient with elevated BP measured during at least one visit, but who did not receive a diagnosis of hypertension or antihypertensive treatment. Elevated BP was defined as a systolic BP equal to or greater than 140 mmHg and/or a diastolic BP of 90 mmHg or greater, in accordance with WHO and the International Hypertension Society (IHS) guidelines for the diagnosis of hypertension.[18] Patients were noted as a diabetes case if they received a diagnosis of diabetes and/or medication to manage diabetes during at least one visit.

Describing antihypertensive treatment and Blood Pressure control

We determined counts of different antihypertensive treatments from both preset medication options and free text entries, classifying each agent by pharmacologic class. We also noted the frequency of different combinations of medications used. Medication use was also broken down by the number of individual patients prescribed each medication regimen at least once (as many patients were seen multiple times during the study period and their regimens frequently changed from visit to visit). We defined BP control as systolic BP <140 mmHg and reported this for confirmed hypertension cases as a percentage of visits for these patients.

Statistical approach for hypertension risk factor analysis

Multivariable linear mixed-effects models were used to analyze the relationship between BP and common hypertension risk factors in the rural communities of Guatemala. We utilized mixed-effects models to account for repeated measurements for many subjects. The primary outcomes were systolic and diastolic pressures. Risk factors included history of hypertension, history of diabetes, age, sex, and/or weight (lb). Continuous variables were centered and scaled while intra-subject correlation was accommodated with random intercepts. Beta estimates and their 95% Confidence Intervals (CI) were reported for the fixed effects.

We also implemented multivariable logistic mixed-effects models to analyze the relationship between elevated BP and confirmed hypertension, respectively, and common risk factors in this rural Guatemalan population. Elevated BP and confirmed hypertension, as defined in the previous section, served as primary outcomes. Risk factors included sex, age, history of diabetes, and/or weight (lb). Body Mass Index (BMI) was not included because height data was frequently missing. Thus, we included weight as a proxy for adiposity/body habitus. Continuous variables were scaled and centered while intra-subject correlation was accommodated using random intercepts. Odds Ratios (OR) and their 95% confidence intervals were reported for the fixed effects.

Multiple Imputation Chain Equations (MICE) were implemented for all analyses because 33% of patient records lacked weight data.[19] As this model assumes that data is missing at random, records with outlier or likely erroneous weights were excluded from analysis rather than imputed. A p-value of 0.05 was considered statistically significant. R (v.4.1.1; R Core Team 2021) and package "Ime4" (v. 1.1.27.1)[20] were used for analysis. In addition, "mice" (v. 3.14.0)[21] and "miceadds" (v. 3.16-18)[22] were used to implement MICE on missing data. Lastly, "lmerTest" (v. 3.1.3)[23] was used to calculate p-values using Satterthwaite approximation. Package "broom.mixed" (v. 0.2.9.3)[24] was used to estimate confidence intervals for the logistic mixed-effects models.

Results

Calculation of hypertension prevalence

We initially screened 6577 patient visits. Of these, we excluded 2931 visits for patients <18 years of age (or with age missing), leaving 3646 visits for adult patients. 3505 of these encounters had a listed diagnosis, representing 2496 unique adult patients, which served as the denominator for calculations of prevalence. Table 1 summarizes the demographic and clinical characteristics of this cohort. We calculated the prevalence of confirmed hypertension, as determined by physician diagnosis or prescription of antihypertensive medication, to be 5.7% (n=142). Additionally, another 11.0% (n=274) of patients had possible hypertension based on BP measurements.

Antihypertensive treatment and Blood Pressure control

Antihypertensive treatments were prescribed in a total of 146 visits, corresponding to 122 unique patients who received at least one antihypertensive prescription at any of their visits. Table 2 provides a breakdown of antihypertensive prescriptions. Angiotensin-Converting Enzyme Inhibitors (ACEIs) were the most commonly prescribed antihypertensive class, representing 59.8% of all pre-

Table 1. Characteristics of patients included in the study sample.

Variables	Overall (n=2496)*
Age (years) (mean (SD))	41.42 (16.77)**
Sex = Male (%)	525 (21.0)
Systolic Blood Pressure (mmHg) (mean (SD))	118.36 (18.04)***
Diastolic Blood Pressure (mmHg) (mean (SD))	74.69 (11.0)***
Patients with confirmed hypertension (%)	142 (5.7)
Patients with elevated Blood Pressure without confirmed hypertension	274 (11.0)
Patients with diabetes	155 (6.2)

SD, Standard Deviation; *Represents unique adult patients with at least one consult with a listed diagnosis; **Two patients with missing age data and 4 with age ≥90 were not included in this calculation; For patients with multiple visits, the mean age for all their visits was factored into this calculation; ***For patients with multiple visits, mean blood pressure for all their visits was factored into these calculations; N=2688, accounting for vitals removed from analysis due to likely erroneous input.

scriptions, followed by a thiazide diuretic, Hydrochlorothiazide (HCTZ), with 28.7% of prescriptions and Calcium Channel Blockers (CCBs) as a distant third (7.3%). Most (89.7%) prescriptions were for single antihypertensives. The most commonly prescribed combination of antihypertensives was an ACEI and HCTZ (Table 3). Of those patients with confirmed hypertension, BP was well-controlled (systolic BP<140 mmHg) at 31.0% of visits.

Analysis of risk factors for elevated Blood Pressure or hypertension diagnosis

We included data from 2688 patient visits in the regression analyses. Of the 2911 visits where any vitals were measured, we removed 210 of these from this analysis due to unsuitable BP measurements and 13 for unsuitable weight measurements. Tables 4 and 5 detail the estimated regression coefficients for the systolic and diastolic BP models, respectively. Male sex, increasing age, increasing weight, and history of hypertension were significantly associated with increasing systolic BP, while history of hypertension, history of diabetes, and increasing weight were significantly associated with increasing diastolic BP. None of the examined independent variables were significantly associated with having at least one elevated BP measurement (≥140/90) or having established hypertension (as defined by a diagnosis of hypertension or receiving antihypertensive treatment). Tables 6 and 7 display odds ratio estimates for these analyses.

Discussion

In this EHR-based study of 2496 Indigenous adults in rural Guatemala, we estimated hypertension prevalence at 16.7% (including confirmed and possible cases). Using EHR data in this manner can allow cost and time savings, while also providing valuable epidemiological data not available in other types of studies.

In this population, systolic BP was significantly higher in

Table 2. Antihypertensive medications prescribed by class and individual medication.

Antihypertensive class	Number of prescriptions (% total)	Medications prescribed (n)
ACEI	98 (59.8)	Enalapril (71), lisinopril (27)
Thiazide diuretic	47 (28.7)	Hydrochlorothiazide (47)
CCB	12 (7.3)	Amlodipine (11), verapamil (1)
BB	5 (3.1)	Metoprolol (5)
ARB	2 (1.2)	Candesartan (1), losartan (1)

ACEI, Angiotensin-Converting Enzyme Inhibitors; CCB, Calcium Channel Blockers; BB, Beta Blockers; ARB, Angiotensin Receptor Blockers.

Table 3. Prescribed antihypertensive regimens arranged by frequency.

Drug regimen	Number of visits (number of unique patients)
ACEI (Enalapril, Lisinopril)	83 (74)
HCTZ	34 (33)
ACEI + HCTZ	10 (10)
Amlodipine	7 (7)
Beta blocker (metoprolol)	5 (5)
ACEI + amlodipine + HCTZ	2 (2)
ACEI + amlodipine	2 (2)
Verapamil	1
ARB (losartan)	1
ACEI + ARB + HCTZ	1

ACEI, Angiotensin-Converting Enzyme Inhibitors; CCB, Calcium Channel Blockers; BB, Beta Blockers; ARB, Angiotensin Receptor Blockers; HCTZ, Hydrochlorothiazide.

Table 4. Regression coefficients for systolic blood pressure model.

Variable	Beta estimate (95% CI)	p
Intercept	115.33 (114.31, 116.34)*	<0.001***
Sex - Male	2.54 (0.86, 4.23)*	0.003**
Age (5 years)	1.44 (1.11, 1.77)*	<0.001***
History of hypertension	26.02 (23.13, 28.91)*	<0.001***
History of diabetes	1.62 (-0.73, 3.98)	0.176
Weight (10 lb)	1.06 (0.42, 1.71)*	0.006**

CI, Confidence Interval.

males and those with a history of diabetes and increased with older age and increasing weight. Diastolic BP was significantly higher in patients with a history of hypertension or diabetes and also increased with increasing weight. However, there were no statistically significant relationships between assessed demographic and anthropometric factors and either confirmed hypertension or elevated BP. Of patients who were prescribed antihypertensive medication, a large majority were treated with a single agent alone. ACEIs were by far the most commonly prescribed antihypertensives, followed by hydrochlorothiazide and CCBs. BP control for patients with confirmed hypertension was poor, with less than one-third of these patients having a systolic BP<140 mmHg.

While our focus was to validate the use of EHR data as a tool, this study also adds to the limited published literature on the epidemiology of hypertension in Guatemala. Our estimate of the prevalence of hypertension in a rural Indigenous population is on par with other studies of similar populations in Guatemala, as well as nationally representative samples. In 2015, Orellana-Barrios and colleagues sampled 1104 adults in the same department (first-level administrative division) from where our data was collected, finding a hypertension prevalence of 12.5%.[25] Chen *et al*. sampled 350 Indigenous adults in 2013, also in this same department, estimating a hypertension prevalence of 18.3%.[26] Steinbrook and colleagues surveyed 806 adults from 2018 to 2019 in two rural, Indigenous municipalities in different departments, estimating hypertension prevalence of 20.3%.[27] Among a nationally representative sample of 1182 women aged 15-49 years of age in Guatemala, 12.7% met the criteria for hypertension in 2017.[28] Our estimated hypertension prevalence of 16.7% falls well within the range of these previously established estimates. Notably, while these previous studies utilized standard epidemiologic sampling to survey their target populations, which can be labor and resource-intensive, we came to a similar estimate using EHR data collected in the course of routine clinical care.

We found that male sex and increasing age and weight (which we used as a proxy for BMI, given that most patients did not have a height recorded) were positively correlated with systolic BP and history of diabetes and weight were positively correlated with diastolic BP. While comparisons are imperfect given the differences in examined variables, previous studies of Guatemalan populations have also found associations between male sex, increasing age, increased adiposity or BMI, history of diabetes, and increasing BP and/or hypertension. Orellana-Barrios and colleagues demonstrated significantly higher BP in men compared to women and significant associations between hypertension and age ≥55 years and elevated waist circumference, but not increased BMI.[25] In their national study of hypertension in women, Pickens *et al.* found that hypertension prevalence increased with increasing age and was more common in those who were overweight or obese, had elevated waist-to-height ratio, and/or had diabetes.[28] Steinbrook *et al.* found an increasing prevalence of hypertension with increasing age.[27] Finally, a multinational study including an urban Guatemalan population found that hypertension was more prevalent in subjects with diabetes and those with elevated BMI or waist

Table 5. Regression coefficients for diastolic blood pressure model.

Variable	Beta estimate (95% CI)	p
Intercept	73.46 (72.75, 74.17)*	<0.001***
Sex - Male	0.66 (-0.49, 1.8)	0.261
Age (5 years)	0.18 (-0.08, 0.43)	0.154
History of hypertension	11.19 (9.03, 13.35)*	<0.001***
History of diabetes	1.82 (0.22, 3.42)*	0.026*
Weight (10 lb)	0.83 (0.47, 1.2)*	0.001**

CI, Confidence Interval.

Table 6. Odds Ratio (OR) estimates for elevated blood pressure model.

Variable	OR (95% CI)	p
Sex - Male	0.82 (0.11, 6.1)	0.844
Age (5 years) (scaled)	1.27 (1, 1.62)	0.053
History of diabetes	1.22 (0.06, 25.22)	0.898
Weight (10 lb) (scaled)	1.12 (0.83, 1.51)	0.371

CI, Confidence Interval.

Table 7. Odds Ratio (OR) estimates for established hypertension model.

Variable	OR (95% CI)	p
Sex - Male	0.64 (0.05, 8.69)	0.735
Age (5 years) (scaled)	1.3 (0.94, 1.8)	0.108
History of diabetes	2.51 (0.2, 31.33)	0.474
Weight (10 lb) (scaled)	1.06 (0.71, 1.57)	0.775

CI, Confidence Interval.

circumference and that those with hypertension were significantly older than those without.[29] Despite the correlations we found, which aligned with the aforementioned literature, none of the examined patient factors displayed statistically significant associations with diagnosed hypertension or an elevated BP measurement (≥140/90 mmHg). However, this is not unexpected, given that dichotomizing continuous outcomes, such as BP, predictably reduce power in statistical analyses.[30]

Most previous studies of hypertension in Guatemala have not included information on antihypertensive treatment and BP control among patients with hypertension. Our use of EHR data allowed us to report on these metrics, which are essential for health systems planning, representing a strength of our study's use of clinical data. While we could not find analogous data in the published literature for rural populations in Guatemala, Quintana and colleagues analyzed data collected from 3246 adults in Guatemala City and found that among those with established hypertension and were prescribed medications, 68.1% were receiving monotherapy, 24.6% two medications, and 7.3% three or more antihypertensives.[31] Of all patients with hypertension, 43.6% had controlled BP. In our sample, a greater proportion of patients with confirmed hypertension who were prescribed medications were prescribed monotherapy (89.7%) and overall BP control was worse (31.0%). We theorize that the greater percentage of monotherapy and worse control in our population compared to the Guatemala City cohort relates to urban/rural differences in regular access to medical care and prescription medication.[32,33] Furthermore, the fact 11.0% (n=274) of patients with BP diagnostic for hypertension did not receive a diagnosis of hypertension or a prescription for antihypertensive medication suggests that hypertension is underdiagnosed and undertreated in these rural communities.

The primary limitations of this study relate to the assessment of hypertension risk factors. Because we utilized EHR-based data rather than an epidemiological survey approach, data for certain risk factors, particularly diabetes status, may have been incomplete. In addition, most records lacked height measurements and thus we had to rely on weight as a proxy for BMI/adiposity. Assuming that most patients were close to average height for this population, weight would correlate closely with BMI (as has been found in previous studies across diverse populations).[34] Prior studies have shown that height in Guatemalan adults is narrowly distributed, with a standard deviation which is less than or equal to that of other global populations.[35-37] However, if there were substantial variations in heights within our sample, this would affect the correlation between BMI and weight. Nevertheless, our analysis of hypertension risk factors was generally in line with previous studies in Guatemalan populations, suggesting that BP is generally higher in males and patients with diabetes and increases with age and not just weight but increased adiposity/body habitus.

The use of EHR data may have also biased the calculation of hypertension prevalence in several ways. First, the measurement of BP and diagnosis and treatment of hypertension was carried out by many different personnel across several medical teams with diverse training and practice backgrounds and using varying equipment. Thus, unlike in epidemiological surveys, these procedures did not follow standard protocols. Therefore, BP and hypertension diagnosis may have been under- or overestimated from one clinic to the next, or even within clinics depending on who measured BP or evaluated patients. Secondly, this data does not represent a random or epidemiologically designed sample. Adults who chose to attend mobile clinics may differ from those who did not attend, in ways that could impact hypertension risk. However, the similarity of our estimates to previously published epidemiologic studies suggests that the ability to include a large number of patients in our analysis helped to minimize the individual impact of these potential biases.

Our experience implementing an EHR in rural Guatemala shows that these efforts do not only improve the delivery and quality of care; as this study illustrates, EHR use can facilitate the evaluation of population health outcomes and, in turn, the planning of future health interventions and quality improvement. A concrete example of this is our team's ongoing NIH-funded program to design, implement, and systematically evaluate a mHealth intervention for hypertension management in these communities.[38] The EHR data collected and analyzed for this study was pivotal to the development of our successful grant proposal. Our experience has implications for similar low-resource settings in LMIC, as the EHR we developed and implemented has a small technological footprint, using readily affordable mobile technology and requiring only intermittent internet access. Thus, the data collected from EHRs like ours can help improve disease surveillance across many settings worldwide.

Conclusions

Using EHR-based data, we estimate a hypertension prevalence of 16.7% in a rural, Indigenous population in Guatemala; almost two-thirds of patients with likely hypertension were not formally diagnosed or treated. Older patients, men, those with diabetes, and those with higher body weights tended to have higher BP. Hypertension treatment was suboptimal in this population, with less than a third of confirmed hypertension patients having good control of their BP. Using EHR-based data is a feasible alternative to traditional epidemiological surveys in low-resource settings, helping to overcome information gaps facing healthcare decision-makers and empowering interventions to address hypertension and other priority chronic diseases.

References

1. World Health Organization (WHO). Noncommunicable diseases. World Health Organization. 2022. Available from: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases
2. World Health Organization (WHO). Hypertension Fact Sheet. 2019. Available from: https://www.who.int/news-room/fact-sheets/detail/hypertension
3. Mills KT, Bundy JD, Kelly TN, et al. Global disparities of hypertension prevalence and control: a systematic analysis of population-based studies from 90 countries. Circulation 2016;134:441-50.
4. Zhang D, Wang G, Joo H. A systematic review of economic evidence on community hypertension interventions. Am J Prev Med 2017;53:S121-30.
5. Hunter DJ, Reddy KS. Noncommunicable diseases. N Engl J Med 2013;369:1336-43.
6. World Health Organization (WHO). Global action plan for the prevention and control of Non-Communicable Diseases. 2021. Available from: http://apps.who.int/iris/bitstream/handle/10665/94384/9789241506236_eng.pdf;jsessionid=171A0B36347E1C6675075C85D01E8A86?sequence=1
7. Checkley W, Ghannem H, Irazola V, et al. Management of NCD in Low- and Middle-Income Countries. Glob Heart 2014;9:431-43.

8. Schutte AE, Srinivasapura Venkateshmurthy N, Mohan S, Prabhakaran D. Hypertension in Low- and Middle-Income Countries. Circ Res 2021;128:808-26.
9. Yegros-Yegros A, van de Klippe W, Abad-Garcia MF, Rafols I. Exploring why global health needs are unmet by research efforts: the potential influences of geography, industry and publication incentives. Health Res Policy Syst 2020;18:47.
10. Charani E, Abimbola S, Pai M, et al. Funders: the missing link in equitable global health research? PLOS Glob Public Health 2022;2:e0000583.
11. Woldemariam MT, Jimma W. Adoption of Electronic Health Record systems to enhance the quality of healthcare in low-income countries: a systematic review. BMJ Health Care Inform 2023;30:e100704.
12. Hung YW, Hoxha K, Irwin BR, et al. Using routine health information data for research in Low- and Middle-Income Countries: a systematic review. BMC Health Serv Res 2020;20:790.
13. World Health Organization (WHO). Non-communicable Diseases (NCD) country profiles: Guatemala. 2018; Available from: https://www.who.int/nmh/countries/gtm_en.pdf?ua=1
14. Instituto Nacional de Estadística Guatemala. Principales Resultados Censo 2018. 2019. Available from: https://www.ine.gob.gt/sistema/uploads/2021/11/19/20211119 2139096rGNQ5SfAlepmPGfYTovW9MF6X2turyT.pdf
15. Internet-in-a-Box. Available from: https://internet-in-a-box.org/
16. Duffy S, Norton D, Kelly M, et al. Using community health workers and a smartphone application to improve diabetes control in rural Guatemala. Glob Health Sci Pract 2020;8:699-720.
17. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. PLoS Med 2007;4:e297.
18. Unger T, Borghi C, Charchar F, et al. 2020 International Society of Hypertension Global Hypertension Practice Guidelines. Hypertension 2020;75:1334-57.
19. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? Multiple imputation by chained equations. Int J Methods Psychiatr Res 2011;20:40-9.
20. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models Usinglme4. J Stat Softw 2015;67:1-48.
21. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations. J Stat Softw 2011;45:1-67.
22. Robitzsch A, Grund S. Miceadds: Some additional multiple imputation functions, especially for “mice”. 2023. Available from: https://CRAN.R-project.org/package=miceadds
23. Kuznetsova A, Brockhoff PB, Christensen RHB. LmerTest package: tests in linear mixed effects models. J Stat Softw 2017;82:1-26.
24. Bolker B, Robinson D, Menne D, et al. broom.mixed: tidying methods for mixed models. 2022. Available from: https://cran.r-project.org/web/packages/broom.mixed/ index.html
25. Orellana-Barrios MA, Nuggent KM, Sanchez-Barrientos H, Lopez-Gutierrez JR. Prevalence of hypertension and associated anthropometric risk factors in indigenous adults of Guatemala. J Prim Care Community Health 2015;6:16-20.
26. Chen D, Rivera-Andrade Á, González J, et al. Prevalence of risk factors for Non-Communicable Diseases in an indigenous community in Santiago Atitlán, Guatemala. Rev Panam Salud Publica 2017;41:e7.
27. Steinbrook E, Flood D, Barnoya J, et al. Prevalence of hypertension, diabetes, and other cardiovascular disease risk factors in two Indigenous municipalities in rural Guatemala: a population-representative survey. Glob Heart 2022;17:82.
28. Pickens CM, Flores-Ayala R, Addo OY, et al. prevalence and predictors of high blood pressure among women of reproductive age and children aged 10 to 14 years in Guatemala. Prev Chronic Dis 2020;17:E66.
29. Organización Panamericana de la Salud Iniciativa Centroamericana de Diabetes (CAMDI): Encuesta de diabetes, hipertensión y factores de riesgo de enfermedades crónicas. 2009. Available from: https://www3.paho.org/hq/dmdocuments/2012/PAHO-CAMDI-Espanol1-2012.pdf
30. Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ 2006;332:1080.
31. Quintana FSW, Casasola MAV, Ortiz Lopez AC, et al. May measurement month 2017-2019: an analysis of blood pressure screening results from Guatemala. Eur Heart J Suppl 2022;24:F16-8.
32. Ippolito M, Chary A, Daniel M, et al. Expectations of health care quality among rural Maya villagers in Sololá Department, Guatemala: a qualitative analysis. Int J Equity Health 2017;16:51.
33. Lawton AM. The right to health in indigenous Guatemala: prevailing historical structures in the context of health care. 2015. Available from: https://www.hhrjournal.org/2015/08/the-right-to-health-in-indigenous-guatemala-prevailing-historical-structures-in-the-context-of-health-care/
34. Diverse Populations Collaborative Group. Weight-height relationships and Body Mass Index: some observations from the Diverse Populations Collaboration. Am J Phys Anthropol 2005;128:220-9.
35. Oyesiku L, Solomons NW, Doak CM, Vossenaar M. Highland Guatemalan women are extremely short of stature, and no lactation duration effects on body composition are observed in a cross-sectional survey. Nutr Res 2013;33:87-94.
36. Subramanian SV, Özaltin E, Finlay JE. Height of nations: a socioeconomic analysis of cohort differences and patterns among women in 54 low- to middle-income countries. PLoS One 2011;6:e18962.
37. Roser M, Appel C, Ritchie H. Human Height. Our World in Data. 2013. Available from: https://ourworldindata.org/human-height
38. Duffy S. mHealth to enable task sharing for hypertension care in LMIC. 2021. Available from: https://reporter.nih.gov/project-details/10471363