

ARTICLE



Predicting the onset of hypertension for workers: does including work characteristics improve risk predictive accuracy?

Tongtan Chantarat ^{1,2}✉, Patricia M. McGovern³, Eva A. Enns¹ and Rachel R. Hardeman^{1,2}

© The Author(s), under exclusive licence to Springer Nature Limited 2022

Despite extensive evidence of work as a key social determinant of hypertension, risk prediction equations incorporating this information are lacking. Such limitations hinder clinicians' ability to tailor patient care and comprehensively address hypertension risk factors. This study examined whether including work characteristics in hypertension risk equations improves their predictive accuracy. Using occupation ratings from the Occupational Information Network database, we measured job demand, job control, and supportiveness of supervisors and coworkers for occupations in the United States economy. We linked these occupation-based measures with the employment status and health data of participants in the Coronary Artery Risk Development in Young Adults (CARDIA) study. We fit logistic regression equations to estimate the probability of hypertension onset in five years among CARDIA participants with and without variables reflecting work characteristics. Based on the Harrell's c- and Hosmer–Lemeshow's goodness-of-fit statistics, we found that our logistic regression models that include work characteristics predict hypertension onset more accurately than those that do not incorporate these variables. We also found that the models that rely on occupation-based measures predict hypertension onset more accurately for White than Black participants, even after accounting for a sample size difference. Including other aspects of work, such as workers' experience in the workplace, and other social determinants of health in risk equations may eliminate this discrepancy. Overall, our study showed that clinicians should examine workers' work-related characteristics to tailor hypertension care plans appropriately.

Journal of Human Hypertension (2023) 37:220–226; <https://doi.org/10.1038/s41371-022-00666-0>

INTRODUCTION

Predicting hypertension onset: the current practice

Risk equations are essential tools that guide clinicians' decisions about which health services to use or preventative strategies to recommend when caring for patients. Knowing patients' risk scores or estimated likelihoods of disease onset also helps clinicians communicate with their patients about the importance and, in some cases, the urgency of behavioral modification to delay disease onset or minimize further progression [1]. Clinicians can use risk equations routinely, leveraging data from various sources. At least 48 different hypertension risk equations have been developed to date [1]. The utility of risk equations and their likelihood of adoption depends not only on their predictive accuracy but also their usability. In general, predictive accuracy means two things: (1) the degree to which risk equations can accurately distinguish between people who will have or not have a disease (i.e., a measure of discrimination) and (2) whether they can predict disease for people at different levels of risk (i.e., a measure of calibration) [2]. Input variables for risk equations should also be readily available to clinicians (e.g., information collected during a regular office visit and stored in medical records). Risk equations that require extensive input variables may be used less given a fast turnaround of each visit [1]. The vast majority of the existing hypertension risk equations predict

disease onset based on age, sex, current blood pressure (often systolic blood pressure or prehypertensive/high-normal status), body mass index (BMI), smoking status, exercise, and family history; [1] in other words, demographic, behavioral, and non-invasive biomarker data.

Another characteristic of risk equations that clinicians often overlook is whether the equations predict a disease equally well for people from different social groups. In recent years, the lack of attention to this characteristic has been highlighted by several empirical studies demonstrating that the disease status of Black patients can be systematically misclassified if a race variable (i.e., a socially constructed characteristic) [3] is embedded in risk equations to adjust for unexplained risk differences between Black and White patients [4]. Using a race-correction term to predict patients' health trajectory without considering health-harming effects of racism and differences in social needs and social determinants of health deprives unhealthy Black patients from access to entitled health services and compensation [5, 6]. This evidence of inequitable care led to calls to remove a race-correction term from risk equations and other aspects of clinical practices [4, 7].

Work as a social determinant of hypertension

The work we do affects our risk of hypertension. Three occupational determinants of hypertension have been extensively

¹Division of Health Policy and Management, University of Minnesota School of Public Health, Minneapolis, MN, USA. ²Center for Antiracism Research for Health Equity, University of Minnesota School of Public Health, Minneapolis, MN, USA. ³Division of Environmental Health Sciences, University of Minnesota School of Public Health, Minneapolis, MN, USA. ✉email: chant083@umn.edu

Received: 29 April 2021 Revised: 8 February 2022 Accepted: 22 February 2022

Published online: 11 March 2022

researched in the past few decades: job insecurity, job loss, and the psychosocial work environment. Job insecurity and job loss, the level of which depends on workers' occupations [8], increase the risk of hypertension among workers directly as added sources of stress and indirectly by disrupting access to essential health-promoting resources (e.g., income, healthcare coverage). Studies from the United States (US) and Scandinavian countries show that workers who have experienced job insecurity or job loss are more likely to be hypertensive than those who have not [9–11]. The degree to which these risk factors affect workers' blood pressure varies based on the workers' gender, age when job loss occurs, and the duration of unemployment [9–11].

While workers' ability to maintain employment is an important predictor of hypertension, the nature of work and the workplace also predict the onset. Well-documented evidence shows that exposure to work-related psychosocial stressors, also referred to as psychosocial work environment (PWE), predicts high stress and increases the risk of hypertension [12]. Three dimensions of PWE predict hypertension onset among workers: job demand, job control, and work-related support [13, 14]. Job demand is characterized by the worker's workload, time pressure, and role conflict. Job control is characterized by the worker's ability to control his/her work activities. This dimension of the PWE consists of two subcomponents: 1) skill discretion (levels of skill and creativity required on the job and the flexibility an employee is permitted in deciding what skills to use); and 2) and decision authority (opportunities for workers to make decisions about their work). The third PWE dimension, work-related support, is characterized by levels of social interaction and support received from coworkers and supervisors. Workers whose occupation is associated with high job demand, low job control, and have limited support face an increased risk of hypertension onset compared to workers whose occupation is low in job demand, high in job control, and receive extensive support [13, 14].

Current study

Despite strong evidence that work characteristics predict the onset of hypertension, no existing risk equations account for these variables. This oversight may represent a missed opportunity for clinicians to better identify workers at high risk of hypertension onset. The lack of risk equations that account for work characteristics also hinders clinicians' ability to look beyond their clinic walls to address a root cause of hypertension inequities in the community they serve. The objective of our study is to examine the extent to which including job loss/job insecurity (measured by employment status) and PWE (measured by the occupation-based measures derived from the publicly available occupation rating data) improve the accuracy of the prediction of hypertension onset for workers. We purposefully excluded a race-correction term from our risk equations and evaluated social determinants of health that are the root cause of racial inequity in hypertension instead.

MATERIALS AND METHODS

Data

Our analysis used the public data from the Coronary Artery Risk Development in Young Adults (CARDIA)—a multi-center prospective study designed to assess exposure to cardiovascular risk factors in young adulthood and patterns of health outcomes in later life. Detailed descriptions of the study design and data collection procedures are published elsewhere [15]. In this secondary analysis, we used participants' work characteristics and relevant hypertension risk factors data from the follow-up assessment conducted in 2000 (Year 15) to predict their hypertension onset in 2005 (Year 20).

CARDIA collected data on a host of cardiovascular risk factors. At all assessments, certified technicians measured the blood pressure of the participants. The systolic (SBP) and diastolic blood pressure (DBP) were measured three times from the participant's right arm using the Hawksley

random-zero sphygmomanometer (Year 15) and the Omron HEM907XL machine (Year 20). We examined non-work-related risk factors as suggested by the current literature [16]. Primary risk factors consisted of each participant's age, gender (male; female), educational attainment (less than high school, some college to having a college degree; higher than college), parental history of hypertension, BMI, and diagnoses of health conditions that are positively associated with hypertension, including high cholesterol, diabetes (type 1 or type 2), heart problems (e.g., heart attack, angina, rheumatic heart disease, mitral valve prolapse), and kidney problems (e.g., urine infection from kidney, kidney stone, kidney problems like nephritis or glomerulonephritis, kidney failure, dialysis, or a kidney transplant). We also included participants' insurance coverage at the Year 15 assessment, as individuals who have difficulty accessing healthcare when needed are less likely to be in control of blood pressure and other conditions co-occurring with hypertension (e.g., diabetes) [17]. Lastly, we also included three health behaviors that are associated with hypertension: [16] drinking status (current drinker; not current drinker), smoking (never smoker, former smoker, current smoker), and physical inactivity during the past year compared to people with the same gender and age (physically inactive; physically active).

As for work-related characteristics, CARDIA assigned participants the 3-digit 1990 US Census Occupational Classification (OCC) based on their response to questions about their industry of employment, occupation, and a class of worker. To determine employment status, participants were asked to describe their current main daily activities and/or responsibilities, with possible responses including working full-time, working part-time, being unemployed/laid off, looking for work, and housekeeping or raising children full time. To assess PWE for workers in our study, we derived the occupation-based measures using the occupation rating data from the Occupational Information Network (O*Net) database [18]. This publicly available online database contains ratings of over 200 occupational characteristics from approximately 900 occupations in the US economy. Ratings of occupational tasks, work activities, knowledge, education and training, work styles, and work context are based on questionnaires completed by a representative sample of current job incumbents. Trained occupational analysts also provide ratings on skills required for particular occupations and how important all rated characteristics are to the success of specific occupations. Since 2005, the O*Net data have been used in more than 60 published studies to examine the extent to which various work-related exposures predict workers' health status. We used similar O*Net items as in previous studies to derive the job demand, job control, and support measures [19, 20]. The possible range is 0–4 for the job demand and job control measures and 0–2 for the support measure, with higher scores indicating greater extent. For a description of the development process of our PWE measures, the values associated with specific occupations, and how we mapped these measures to the CARDIA data, please refer to the online supplements S1–S3.

Exclusion criteria

The total number of participants who received an assessment in Year 15 was 3,671. We restricted our analytical sample to participants with non-military occupations [21] who reported working full-time, part-time, and unemployed, laid off, or actively looking for jobs in Year 15. Participants were excluded if they worked in occupations to which the PWE measures could not be assigned (see online supplement S4), if they were hypertensive or had a history of hypertension before Year 15, were pregnant at the time of the assessment (women only), had a missing hypertension status in Year 20, or had missing covariates described above. Figure 1 displays the number of excluded participants in each exclusion step to arrive at the final analytical sample ($N = 1,909$). The age, gender, and educational makeup of our final analytical sample were not significantly different from the excluded group. However, the proportion of Black participants in our analytical sample was significantly smaller than those in the excluded group (37.4% vs. 54.3%, $p < 0.001$).

Model estimation and validation

We defined hypertension as having SBP ≥ 140 mmHg, DBP ≥ 90 mmHg, and/or if the participant was taking hypertensive medication, consistent with the Seventh Report of Joint National Committee on Prevention, Detection, Evaluation, and Treatment of Hypertension guideline (JNC-7) [22]. Participants with prehypertension were those with SBP of 120 to 139 mmHg and/or DBP of 80–89 mmHg [22]. We fitted a logistic regression to estimate a five-year probability of hypertension. Similar to other US-based models [23–25], we minimized the use of interaction terms in our models.

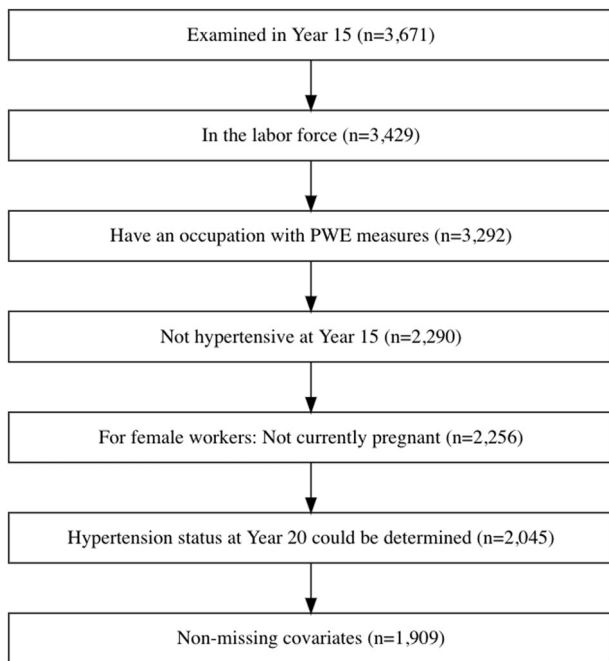


Fig. 1 Sample Exclusion.

We tested 2-way interactions between the three PWE measures with gender and age groups, grounded on the evidence that these demographic characteristics likely affect ones' experience in the workplace [20]. We employed a backward selection algorithm to arrive at the most parsimonious model [26]. For each round of deletion, the algorithm calculated the Bayesian Information Criterion (BIC), which combines the model likelihood with a penalty for the number of predictors in the model; [27] the model with the lowest BIC was considered the most parsimonious model. To ensure that the final model reasonably predicted hypertension among workers and takes into consideration the employment status and PWE, we purposefully programmed the selection algorithm to generate several models: (1) retained variables that were significantly associated with the five-year probability of hypertension (simplified model); (2) retained significant variables *and* employment status and all three dimensions of PWE regardless of their *p* value (full PWE model); and (3) retained significant variables, employment status, and then only those PWE dimension that were significantly associated with hypertension (partial PWE model). For each PWE model, we also tested whether retaining participants' gender and age on top of the work-related variables further improve the model's performance or not.

We validated our models using two metrics. Model discrimination was evaluated using the Harrell *c*-statistic; predictive models with a *c*-statistic of 0.7 or higher are typically considered to have good discrimination [28]. Model calibration was evaluated using the Hosmer–Lemeshow goodness-of-fit statistic, which is distributed as chi-squared. A statistically significant result for the goodness-of-fit test indicates that there is a statistically significant difference in the predicted number of outcomes compared to the observed number of outcomes for at least one group (i.e., poor calibration).

Comparison of model prediction accuracy for Black and White participants

We compared the *c*-statistics evaluated for the Black participants only and White participants only. Because our analytical sample contained an unequal number of Black ($n = 733$) and White ($n = 1,176$) participants, this sample size difference may contribute to differences in model prediction accuracy between these groups. We, therefore, created 200 datasets of White participants of the same sample size as that of Black participants (i.e., $n = 733$) by sampling without replacement from the original sample of White participants. We then repeated the estimation process described above for the combined Black and White samples (total $n = 1,466$ for each dataset). The mean of the *c*-statistics estimated for White participants based on all 200 samples was compared to the *c*-statistic evaluated using the full sample of Black participants with two-sample T-tests.

All data management and analyses were conducted in R version 3.5.2. The Institutional Review Board of the University of Minnesota reviewed this study protocol and designated it as non-human research.

RESULTS

Table 1 displays the characteristics and hypertension status of participants in our sample. After applying the exclusion criteria, our sample consisted of 1,909 participants with a mean age of 40.3 years that were relatively balanced in sex (54% female and 46% male) but unbalanced in race (38% Black and 62% White). A large majority of the participants in our sample worked full-time (81%) or part-time (12%). The mean job demand, job control, and support scores were 1.8 ($SD = 0.3$), 2.7 ($SD = 0.6$), 1.2 ($SD = 0.2$), respectively. Overall, the participants were generally healthy, with a mean BMI of 27.9 and prevalence of hypertension co-occurring conditions lower than the national average [29]. Most participants reported drinking alcohol, never smoke, and being physically active. At Year 20, 11.3% of participants who were hypertension-free at Year 15 developed hypertension.

Table 2 displays the parsimonious models with the lowest BIC for the five backward selection strategies. When no restriction was imposed on the selection (simplified model), the backward selection identified the logarithm of the BMI, being a current smoker, being prehypertensive, and having a family history of hypertension as predictors of hypertension onset in the next five years. When we purposefully programmed the selection algorithm to retain employment status and all three dimensions of the PWE (full PWE model), all predictors identified in the simplified model remained significant in this model. Levels of job control associated with the worker's occupation was the only PWE dimension that predicted hypertension onset in five years with $p < 0.05$. When participant's gender and age were also retained (full PWE model with gender and age), the coefficients of the predictors changed slightly compared to in the full PWE model. When we purposefully programmed the selection algorithm to retain only employment status and job control (partial PWE model), the final model identified the same set of predictors as in the full PWE model. The addition of gender and age (partial PWE model with gender and age) did not change the effect size of the model predictors significantly. All five models described earlier had high discrimination (greater than 0.8) and were well-calibrated. The full PWE model with gender and age, the partial PWE model, and the partial PWE model with gender and age had the highest *c*-statistics (0.805), but the simpler full PWE predicts hypertension onset almost equally well (*c*-statistic = 0.804) with slightly better calibration (Hosmer–Lemeshow $\chi^2 = 6.09$).

Table 3 displays the mean *c*-statistics of the five models evaluated with 200 datasets with all Black and equal number of sampled White participants. All models predicted hypertension well for both Black and White participants, as indicated by *c*-statistics of greater than 0.7. We observed significantly lower *c*-statistics when using data of Black participants compared to using data of sampled White participants in all models. This finding suggests that even when the prediction models are estimated using the same number of Black and White participants, our models predict hypertension less accurately for Black participants.

DISCUSSION

Our study investigated whether including employment status (as a proxy of job loss and instability) and psychosocial work environment in risk equations improves the accuracy of prediction of the short-term onset of hypertension. Despite work-related risk factors' direct and indirect linkages with health and well-being, clinicians have never used such information to identify workers at elevated risk of hypertension. Unlike biomedical indicators (e.g., age, weight, height) or health behaviors routinely assessed during clinical intake

Table 1. Analytical Sample Characteristics at Year 15 and Their Hypertension Status at Year 20 ($n = 1,909$).

| | Frequency | Percent |
|---|-----------|---------|
| Sociodemographic characteristics | | |
| Age (Mean, SD) | 40.3 | (3.6) |
| Black | 733 | (38.4%) |
| Female | 1,031 | (54.0%) |
| Less than high school | 343 | (18.0%) |
| Some college to college degree | 1,070 | (56.1%) |
| Higher than college | 496 | (26.0%) |
| Insurance coverage | | |
| Employer-sponsored | 1,575 | (82.5%) |
| Medicaid/Medicare | 80 | (4.2%) |
| Military insurance | 27 | (1.4%) |
| Self-insured | 211 | (11.1%) |
| Uninsured | 1 | (0.1%) |
| Health conditions | | |
| Body mass index (Mean, SD) | 27.9 | (6.0) |
| High cholesterol | 305 | (16.0%) |
| Heart problem | 201 | (10.5%) |
| Diabetes | 69 | (3.6%) |
| Kidney problem | 112 | (5.9%) |
| Health behaviors | | |
| Drinker | 1,572 | (82.3%) |
| Never smoker | 1,209 | (63.3%) |
| Former smoker | 353 | (18.5%) |
| Current smoker | 347 | (18.2%) |
| Physically inactive | 103 | (5.4%) |
| Employment status | | |
| Work full time | 1,548 | (81.1%) |
| Work part time | 231 | (12.1%) |
| Unemployed/Looking for work | 130 | (6.8%) |
| Psychosocial work environment (Mean, SD) | | |
| Job demand (0–4) | 1.8 | (0.3) |
| Job control (0–4) | 2.7 | (0.6) |
| Job support (0–2) | 1.2 | (0.2) |
| Hypertension history and status | | |
| Have at least one parent with hypertension ^a | 1,290 | (67.6%) |
| Prehypertension - Year 15 ^b | 504 | (26.4%) |
| Hypertension - Year 20 ^c | 215 | (11.3%) |

^aAt least one parent or both parents with history of hypertension

^bPrehypertension: SBP 120–139 mmHg or DBP 80–89 mmHg

^cHypertension: SBP \geq 140 mmHg or DBP \geq 90 mmHg or currently taking hypertension medication

(e.g., smoking, drinking, physical activity), consideration of work characteristics is only possible if a worker volunteers such information to their provider. As clinicians start to look beyond the walls of their clinics and collect data on social determinants of health that may affect their patients' hypertension status [30], our study provides new insight into how this information can help improve hypertension care. We found that risk equations that incorporate workers' biomedical risk factors as well as employment status and PWE predict the onset of hypertension more accurately than ones that use only biomedical risk factors. However, we also found that our enhanced risk equations predict hypertension onset

more accurately for White than Black workers, even when the data used to derive such equation consists of the same number of data points from the two racial groups.

One of the major barriers to using work characteristics to predict disease onset is ensuring that such information is collected appropriately. While asking workers about their employment status may be difficult on its own and may require better data collection strategies or linkage of data from other sources [31], assessing workers' psychosocial work environment during a short clinical encounter can pose another challenge. The risk equations we presented in this paper used the occupation-based measures of job demand, control, and support that can easily be coded from a worker's response about the nature of work they do. In light of the COVID-19 pandemic and the need to better surveil the health status of US workers, the National Institute of Occupational Health and Safety has released a guideline that employers and health professionals can use for occupational coding based on a response to the question "what is your occupation?" [32]. Compared to the 49-question Job Content Questionnaire (JCQ) [33], which is used commonly by occupational health scientists to assess PWE, the implementation of occupation-based PWE measures requires significantly less time and the occupational coding process can be made automated in the electronic medical records system. Institutional initiatives to streamline the collection of workers' work-related data will ensure the effective adoption of risk equations like ours and others that may have been in development for clinicians to predict hypertension trajectory and tailor hypertension care for workers more effectively. Future research should examine benefits and costs of automating the calculation of hypertension risk and making this information readily available for clinicians to guide their decisions.

It is worth noting that occupation-based PWE measures used in our models assess the *average* level of exposure to work-related stressors for a particular occupation, but not necessarily the level that a specific worker experiences [20]. Past research suggests that workers from different racial backgrounds experience different levels of PWE, even when they have the same occupation [20, 34–36]. In a racist society like the US and countries with a colonial past, residential segregation and structural racism in education jointly produce the workforce in which Black workers, particularly Black women, disproportionately hold low-status occupations relative to White workers [5]. However, even Black workers holding high-status occupations experience a more stressful and challenging PWE than their White peers. These Black workers frequently report being tokenized by their White coworkers [36]. They are also more likely to be assigned "diversity work" and experience workplace discrimination and harassment than Black workers with low-status occupations, let alone their White coworkers [35]. Different lived experiences in the workplace between Black and White workers with the same occupation may explain why our risk equations that use occupation-based PWE predict hypertension onset less accurately for Black than White workers.

While our choice to use the occupation-based PWE measures in our hypertension risk equations had the ease of data collection and the likelihood of adoption in the fast-paced clinical setting in mind, our findings of lower prediction accuracy for Black workers signal the need for additional research on ways to utilize work-related information in the real-world clinical setting. Discrepancies in prediction accuracy by workers' race cannot/must not be corrected with the race-correction term; while adding an interaction term between race and PWE would capture the heterogeneous *effect* of PWE on hypertension onset by race, it does not correct the fact that workers from different racial groups have different PWE experiences for a given occupation. Future research should investigate if including other individual-level measures of social conditions, particularly those that have already been captured in electronic health records (e.g., financial resources, living arrangements, social connections) [37] will minimize discrepancies in prediction accuracy and promote

Table 2. Logistic Regression Models Predicting Five-year Hypertension Onset and Their Validation Statistics.

| Predictor | Simplified | | Full PWE | | Full PWE with Gender and Age | | Partial PWE | | Partial PWE with Gender and Age | |
|---|------------|------|----------|------|------------------------------|--------|-------------|-------|---------------------------------|-------|
| | Coeff | SE | Coeff | SE | Coeff | SE | Coeff | SE | Coeff | SE |
| Model estimation | | | | | | | | | | |
| Intercept | -12.24 | 1.29 | *** | 1.71 | *** | 1.99 | *** | 1.38 | *** | 1.68 |
| Log BMI | 2.619 | 0.38 | *** | 0.38 | *** | 2.497 | *** | 2.55 | *** | 2.496 |
| Current smoker | 0.836 | 0.18 | *** | 0.18 | *** | 0.775 | *** | 0.757 | *** | 0.77 |
| Currently prehypertensive ^a | 0.662 | 0.19 | *** | 0.16 | *** | 1.71 | *** | 1.677 | *** | 1.706 |
| Family history of hypertension ^b | 1.667 | 0.16 | *** | 0.19 | *** | 0.656 | *** | 0.668 | *** | 0.65 |
| Work full time | | | -0.077 | 0.31 | | -0.059 | 0.31 | | -0.084 | 0.31 |
| Work part time | | | -0.11 | 0.38 | | -0.135 | 0.38 | | -0.115 | 0.38 |
| Job demand score (0-4) | | | 0.157 | 0.3 | | 0.167 | 0.3 | | -0.115 | 0.38 |
| Job control score (0-4) | | | -0.344 | 0.17 | * | -0.347 | 0.17 | * | -0.301 | 0.13 |
| Support score (0-2) | | | -0.178 | 0.61 | | -0.16 | 0.62 | | -0.306 | 0.13 |
| Age | | | | | | 0.024 | 0.02 | | 0.025 | 0.02 |
| Female | | | | | | 0.197 | 0.17 | | 0.188 | 0.17 |
| Model validation | | | | | | | | | | |
| C-statistic | 0.802 | | 0.804 | | 0.805 | | 0.805 | | 0.805 | |
| Hosmer-Lemeshow χ^2 | 8.449 | | 6.090 | | 9.176 | | 8.539 | | 10.317 | |
| Hosmer-Lemeshow <i>p</i> value | 0.391 | | 0.637 | | 0.328 | | 0.383 | | 0.243 | |

****p* < 0.001; ***p* < 0.01, **p* < 0.05

^aPrehypertension: SBP 120-139 mmHg or DBP 80-89 mmHg

^bAt least one parent or both parents with history of hypertension

Table 3. Comparison of the Mean C-statistics of the 200 Regression Models Estimated Using the Datasets That Consisted of All Black Participants (*n* = 733) and the Same Number of Randomly Sampled White Participants (*n* = 733)[†].

| Model | Race of the participant | Mean C-statistic |
|------------------------------------|-------------------------|------------------|
| Simplified*** | white | 0.818 |
| | Black | 0.762 |
| Full PWE*** | white | 0.817 |
| | Black | 0.761 |
| Full PWE with gender and age*** | white | 0.819 |
| | Black | 0.763 |
| Partial PWE*** | white | 0.819 |
| | Black | 0.760 |
| Partial PWE with gender and age*** | white | 0.815 |
| | Black | 0.763 |

****p* < 0.001; ***p* < 0.01, **p* < 0.05

[†]*p* values were based on the two-sample Student's *t* test comparing the mean c-statistics for the best fitted models estimated from 200 samples.

equity in healthcare for workers of all backgrounds. Additionally, researchers should investigate whether the traditional JCQ can be abbreviated for a clinical intake form and evaluate whether the racial discrepancy in the prediction accuracy by race also exists in risk equations that incorporate the JCQ responses.

Our findings should be considered in light of several limitations. First, because of our relatively small sample size after exclusion, we were unable to estimate and validate our models using two separate datasets. It is possible that the validation statistics we calculated may be overly optimistic. Hence, our risk equation should be validated with data from other longitudinal cardiovascular studies. At the time of writing, this was a challenging task because although most, if not all, longitudinal studies measured participants' employment status, they did not assess participants' occupations in detail, hindering the merging of occupation-based PWE measures with biomedical risk data. Second, our study defined hypertension status based on the JNC-7 guideline. Although this procedure is consistent with the standard of care around the time the study data was collected, the American Heart Association (AHA) and the American College of Cardiology (ACC) released the updated guideline in 2017 that lower the clinical threshold for hypertension to 130/80 mmHg and eliminate the prehypertension category altogether. We encourage future studies with data collected after the release of the AHA/ACC guideline to replicate our protocol to examine the sensitivity of our findings. Lastly, the data used to derive our risk equations came from Black and White workers only. Thus, the risk equations presented in this paper should be examined for their prediction accuracy with data from workers from other racial backgrounds before their adoption in real-world practices.

Work is an important social determinant of hypertension and should be included in the risk prediction equations. This effort will enhance clinicians' ability to make better decisions for hypertension care for workers with varied working conditions, which has the potential to address population health inequities. We demonstrated that including workers' employment status and psychosocial work environment in risk equations improves their prediction accuracy. To maximize the utility of risk equations that include social determinants of health like work conditions, we must also determine how to effectively capture this information in ways that do not disrupt a clinical workflow. It is also critical to pay attention to racial implications, big or small, that can arise when developing decision tools like risk equations and address such issues appropriately.

Summary

What is known about the topic

- Despite extensive evidence that work is an important determinant of hypertension, none of the existing risk equations use patients' work-related characteristics to predict disease onset.
- Several risk equations incorporate a race-correction term. This embedded racist practice has been shown to produce racial inequity in patient care.

What this study adds

- We estimated and compared the prediction accuracy of risk equations with and without variables reflecting work-related characteristics.
- We did not treat race as a hypertension risk predictor. Instead, we fit the "raceless" model and tested if the resulting risk equations predict hypertension equally well for Black and White people.
- We found that the risk equations with work-related variables predict hypertension onset more accurately than those without. Our models predict hypertension onset more accurately for White than Black people.

REFERENCES

1. Sun D, Liu J, Xiao L, Liu Y, Wang Z, Li C, et al. Recent development of risk-prediction models for incident hypertension: an updated systematic review. *PLoS One*. 2017;12:e0187240. <https://doi.org/10.1371/journal.pone.0187240>.
2. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*. 2007;115:928–35. <https://doi.org/10.1161/CIRCULATIONAHA.106.672402>.
3. Smedley A, Smedley BD. Race as biology is fiction, racism as a social problem is real: anthropological and historical perspectives on the social construction of race. *Am Psychol*. 2005;60:16–26. <https://doi.org/10.1037/0003-066X.60.1.16>.
4. Roberts DE. The art of medicine: abolish race correction. *Lancet* 2020;397:17–18. [https://doi.org/10.1016/S0140-6736\(20\)32716-1](https://doi.org/10.1016/S0140-6736(20)32716-1).
5. McClure ES, Vasudevan P, Bailey Z, Patel S, Robinson WR. Racial capitalism within public health—how occupational settings drive Covid-19 disparities. *Am J Epidemiol*. 2020;189:1244–53. <https://doi.org/10.1093/aje/kwaa126>.
6. Kakani P, Chandra A, Mullainathan S, Obermeyer Z. Allocation of COVID-19 relief funding to disproportionately black counties. *JAMA*. 2020;324:1000–3. <https://doi.org/10.1001/jama.2020.15301>.
7. Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight — reconsidering the use of race correction in clinical algorithms. *N Engl J Med*. 2020;383:874–82. <https://doi.org/10.1056/nejmms2004740>.
8. Burgard SA, Brand JE, House JS. Perceived job insecurity and worker health in the United States. *Soc Sci Med*. 2009;69:777–85. <https://doi.org/10.1016/j.socscimed.2009.06.029>.
9. Levenstein S, Smith MW, Kaplan GA. Psychosocial predictors of hypertension in men and women. *Arch Intern Med*. 2001;161:1341–6. <https://doi.org/10.1001/archinte.161.10.1341>.
10. Nygren K, Hammarström A, Gong W. Is hypertension in adult age related to unemployment at a young age? results from the Northern Swedish cohort. *Scand J Public Health*. 2015;43:52–58. <https://doi.org/10.1177/1403494814560845>.
11. Acevedo P, Mora-Urda AI, Montero P. Social inequalities in health: duration of unemployment unevenly effects on the health of men and women. *Eur J Public Health*. 2019;30:305–10. <https://doi.org/10.1093/eurpub/ckz180>.
12. Gilbert-Ouimet M, Trudel X, Brisson C, Milot A, Vézina M. Adverse effects of psychosocial work factors on blood pressure: systematic review of studies on demand-control-support and effort-reward imbalance models. *Scand J Work Environ Heal*. 2014;40:109–32. <https://doi.org/10.5271/sjweh.3390>.
13. Karasek RA. Job demand, job decision latitude, and mental strain: implications for job redesign. *Adm Sci Q*. 1979;24:285–308. <https://about.jstor.org/terms>.
14. Johnson JV, Hall EM. Job strain, work place social support, and cardiovascular disease: a cross-sectional study of a random sample of the swedish working population. *Am J Public Health*. 1988;78:1336–42. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1349434/pdf/amjph00249-0078.pdf>.

15. Friedman GD, Cutter GR, Donahue RP, Hughes GH, Hulley SB, Jacobs DR, et al. CARDIA: study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol*. 1988;41:1105–16. https://ac.els-cdn.com/0895435688900807/1-s2.0-0895435688900807-main.pdf?_tid=8b551279-e8e3-434c-a9c7-95b89bb73c03&acdnat=1551994503_ee286aa869d2c7dec373ad5b73138216.
16. Benjamin EJ, Virani SS, Callaway CW, Chamberlain AM, Chang AR, Cheng S, et al. Heart disease and stroke statistics-2018 update a report from the American Heart Association. *Circulation*. 2018;137:67–492. <https://doi.org/10.1161/CIR.0000000000000558>.
17. Kressin NR, Terrin N, Hanchate AD, Price LL, Moreno-Koehler A, LeClair A, et al. Is Insurance Instability Associated with Hypertension Outcomes and Does This Vary by Race/ethnicity? *BMC Health Serv Res*. 2020;20. <https://doi.org/10.1186/s12913-020-05095-8>.
18. O*NET Resource Center. O*NET Resource Center. <https://www.onetcenter.org/>. Published 2019. Accessed July 9, 2019.
19. Cifuentes M, Boyer J, Gore R, D'Errico A, Tessler J, Scollin P, et al. Inter-method agreement between O*NET and survey measures of psychosocial exposure among healthcare industry employees. *Am J Ind Med*. 2007;50:545–53. <https://doi.org/10.1002/ajim.20480>.
20. McCluney CL, Schmitz LL, Hicken MT, Sonnega A. Structural racism in the workplace: does perception matter for health inequalities? *Soc Sci Med*. 2018;199:106–14. <https://doi.org/10.1016/j.socscimed.2017.05.039>.
21. Shrestha A, Ho TE, Vie LL, Labarthe DR, Scheier LM, Lester PB, et al. Comparison of Cardiovascular Health Between US Army and Civilians. *J Am Heart Assoc*. 2019;8. <https://doi.org/10.1161/JAHA.118.009056/FORMAT/EPUB>.
22. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, et al. Seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure. *Hypertension*. 2003;42:1206–52. <https://doi.org/10.1161/01.HYP.0000107251.49515.c2>.
23. Parikh NI, Pencina MJ, Wang TJ, Benjamin EJ, Lanier KJ, Levy D, et al. A risk score for predicting near-term incidence of hypertension: The Framingham Heart Study. *Ann Intern Med*. 2008;148:102. <https://doi.org/10.7326/0003-4819-148-2-200801150-00005>.
24. Paynter NP, Cook NR, Everett BM, Sesso HD, Buring JE, Ridker PM Prediction of Incident Hypertension Risk in Women with Currently Normal Blood Pressure. *Am J Med*. 2009;122:464–71. <https://doi.org/10.1016/j.amjmed.2008.10.034>.
25. Kshirsagar AV, Chiu Y, Bomback AS, August PA, Viera AJ, Colindres RE, et al. A Hypertension Risk Score for Middle-aged and Older Adults. *J Clin Hypertens (Greenwich)*. 2010;12:800–8. <https://doi.org/10.1111/j.1751-7176.2010.00343.x>.
26. Draper N, Smith H. Selecting the "Best" Regression Equation. In: Draper N, Smith H, eds. *Applied Regression Analysis*; 1998. <https://doi.org/10.1002/9781118625590.ch15>.
27. Neath AA, Cavanaugh JE. The Bayesian information criterion: background, derivation, and applications. *Wiley Interdiscip Rev Comput Stat*. 2012;4:199–203. <https://doi.org/10.1002/wics.199>.
28. Hosmer DW, Lemeshow S, Sturdivant RX. *Applied Logistic Regression*. 2nd ed. Wiley-Interscience Publication; 2000.
29. National Institute for Occupational Safety and Health. CDC - NIOSH Worker Health Charts. <https://wwwn.cdc.gov/NIOSH-WHC/chart/nhis-chronic?OU=HYPEV&T=R&V=R>. Published 2019. Accessed December 2, 2019.
30. Daniel H, Bornstein S, Kane G. Addressing Social Determinants to Improve Patient Care and Promote Health Equity: An American College of Physicians Position Paper. *Ann Intern Med*. 2018;168. <https://doi.org/10.2105/AJPH>.
31. Cantor MN, Thorpe L. Integrating Data on Social Determinants of Health into Electronic Health Records. *Health Aff*. 2018;37:585–90. <https://doi.org/10.1377/hlthaff.2017.1252>.
32. Marovich S, Mobley A, Groenewold M. Making Industry and Occupation Information Useful for Public Health: A Guide to Coding Industry and Occupation Text Fields. <https://blogs.cdc.gov/niosh-science-blog/2020/06/17/industry-occup-coding/>. Published 2020. Accessed January 11, 2019.
33. Karasek R, Brisson C, Kawakami N, Houtman I, Bongers P, Amick B. The job content questionnaire (JCQ): an instrument for internationally comparative assessments of psychosocial job characteristics. *J Occup Health Psychol*. 1998;3:322–55. <https://doi.org/10.1037/1076-8998.3.4.322>.
34. Braboy Jackson P, Thoits PA, Taylor HF. Composition of the workplace and psychological well-being: the effects of tokenism on America's Black Elites. *Soc Forces*. 1995;74:543–57. <https://www-jstor-org.ezp3.lib.umn.edu/stable/pdf/2580491.pdf?refreqid=excelsior%3Ab0fa0131c642f3d01322fcca6ecfb747>.
35. Wingfield AH. Flatlining: Race, Work, and Health Care in the New Economy. 1st ed. University of California Press; 2019.
36. Wingfield AH, Wingfield JH. When visibility hurts and helps: how intersections of race and gender shape black professional men's experiences with tokenization. *Cult Divers Ethn Minor Psychol*. 2014;20:483–90. <https://doi.org/10.1037/a0035761>.
37. Chen M, Tan X, Padman R. Social determinants of health in electronic health records and their impact on analysis and risk prediction: a systematic review. *J Am Med Inf Assoc* 2020;27:1764–73. <https://doi.org/10.1093/jamia/ocaa143>.

AUTHOR CONTRIBUTIONS

TC acquired the study funding, designed and conducted all the analyses, and drafted the original manuscript. PMM, EAE, and RRH supervised the design of all the analyses, reviewed, and critically revised the final manuscript. All authors approved the final version of the manuscript for submission.

FUNDING

This research was funded by the Midwest Center for Occupational Health and Safety Pilot Project Research Training Program, supported by the National Institute for Occupational Safety and Health (NIOSH) under award number T42OH008434. The contents of this effort are solely the responsibility of the authors and do not necessarily represent the official views of NIOSH, the Centers for Disease Control and Prevention, or other associated entities.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41371-022-00666-0>.

Correspondence and requests for materials should be addressed to Tongtan Chantarat.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.