Alysen Casaccio

01_Early Programming

Hypertension Disparities in Black Americans

July 23, 2023

Introduction

Hypertension also referred to as high blood pressure, is when the force of blood flowing through a person's vascular system is consistently higher than usual. Left untreated, the damage this does to the circulatory system is a significant contributing factor to heart attack, stroke, and other life-threatening health issues. Most of the time, this condition has no obvious symptoms, but it is the most prevalent¹ and modifiable risk factor for cardiovascular disease, making identification and treatment essential for all populations.

The treatment and control of hypertension is a significant public health challenge, but despite efforts to improve outcomes, there are known to be significant differences in the prevalence, treatment, and management of hypertension, specifically for non-Hispanic Black, Hispanic, and Asian Americans.² In addition, a recent analysis³ found that Black Americans are diagnosed with hypertension earlier in life and experience 4-5 times greater hypertension-related mortality than white Americans, with 30% higher risk of stroke, 50% higher risk of cardiovascular disease mortality, and more than four times higher risk of end-stage renal disease.

Problem Statement

The hypothesis I would like to explore involves the root question of what is causing the underlying disparity in the first place. My problem statement asks, "Are hypertension disparities in Black Americans related to discrimination experiences in the United States?"

A study of discrimination in America, published by Health Services Research⁴ studied the odds of discrimination among Black Americans (via telephone study), to examine the variation found by socioeconomic status, gender, and neighborhood racial composition. Most relevant to my topic, they

asked respondents about their experiences of discrimination in relation to several domains, one of which was discrimination experiences while navigating healthcare.

Their findings reported that about one-third of Black Americans surveyed reported experiencing discrimination in clinical encounters, while 22% avoided seeking healthcare for themselves or family members due to anticipated discrimination. The study also found that among Black Americans, having a college degree was associated with higher odds of experiencing overall institutional discrimination.

The article, "Minorities' Diminished Return Theory" ⁵ discusses this in depth because although it is well known that a higher socioeconomic status carries with it protective effects on population health, there is a growing body of research that shows a smaller health return, specifically, for Black Americans when compared to white Americans. Are

Since neither higher education nor higher income is a strong protective measure against hypertension for Black Americans, like it is for white Americans, this leads us to look for and examine differences in hypertension identification, treatment, and control in the US.

Analysis and Findings

The National Center for Health Statistics has conducted full health and condition-specific surveys of Americans annually since 1957 for the purposes of public health initiatives. The data connected to these surveys is free to access and hosted through the National Health Interview Survey (NHIS) and the Centers for Disease Control and Prevention (CDC), depending on the year the survey was conducted.

For the purposes of this project, I exported two datasets for comparison, from 1974 and 2019, respectively, and in both cases, focused on the general survey called the "person files" and the more specific "hypertension supplement." Within these datasets are many variables, and since the survey questions themselves are revised and updated about every ten years, the combined dataset did require

some amount of adjustment and normalization, which is described in detail in Appendix 1, along with screenshots of the survey questions, possible responses, and mapping decisions made to normalize the data for the following analysis.

After removing data that either was not relevant to my current query or was simply too expansive for the project, we were left with the following set of fourteen variables:

- Survey Year
- Race
- Gender
- Education Level
- Family Income
- Informed You Have Hypertension
- Taking a Medication to Treat Hypertension
- Time Since Last Doctor's Visit
- Height
- Weight
- Body Mass Index (BMI)
- Comorbidity: Diabetes
- Comorbidity: Stroke
- Comorbidity: Cardiovascular Disease

My focus for the problem statement is on non-Hispanic, Black American men, so this is the population I will be considering most closely when comparing and analyzing the data at hand.

The first and most striking finding pointed to the smaller percentage of Black men who were respondents. The total respondent population for both years was 57,562 individuals, and in both 1974 and 2019, Black men represented only 3% and 4% of the totals. When segmenting the population of all Black Americans, the rates of respondents were 9% in 1974 and 11% in 2019.

This is important to examine because if the survey is being used for public health purposes, and Black men are under-represented, then the dataset that we use to propose solutions may have bias, which prevents those solutions from being effective for the underrepresented population. This may be in alignment with the percentage represented in the United States Census, by percentage, however.

See image 1 for a visual segmentation of both race and gender demographics for the respondents in both survey years.

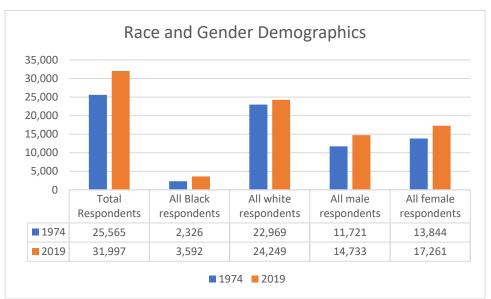


Image 1.1: Race and Gender Demographics for the National Health Interview Survey, 1974 and 2019

The measurements used which surrounded hypertension will focus on three areas: Patients who were told they have hypertension, patients who were prescribed a medication to treat their hypertension, and patients who were taking a prescribed medication and had seen a doctor in the past twelve months. I divided these measurements by survey year to provide a comparison of identification and treatment in 1974 versus today, but I also segmented the responses, comparing the experience white males appeared to have with the experience had by black males, again, focusing on the identification of hypertension, pharmaceutical treatment, and timing of the last reported doctor visits.

The analysis of these basic data points revealed some interesting patterns. First, let us look at these three treatment data points from 1974, segmented by Black or white males. See Image 1.2.

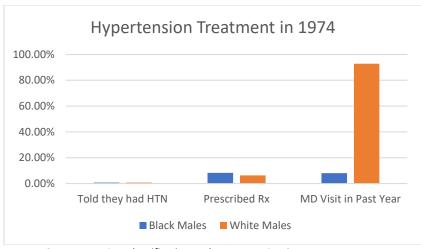


Image 1.2: Hypertension Identification and Treatment in 1974.

I did not want my visualizations to be made unclear by the more significant number of white males who participated in the surveys, so I generated percentage outcome measurements using a simple quality measure format. (Example: Denominator – Total number of Black males prescribed hypertension medication in 1974. Numerator – Total number of the respondents in the denominator who reported having a doctor's visit in the past 12 months. Numerator/Denominator = percentage rounded .02 decimal points.)

One aspect that stands out in the 1974 survey is the number of respondents who had not ever been told they had hypertension, but who were regularly taking hypertension medications in treatment for the condition. The more obvious disparity here, of course, is the significantly larger number of white males who had a doctor's visit in the past year. This third measurement only includes the respondents who already informed us they were actively taking a hypertensive medication, so the lack of doctor-patient relationship in Black males may undoubtedly have played a role in health outcomes at that time.

Looking now at the 2019 data in Image 1.3, we can see that Black males who are taking hypertensive medication are more likely to have had a doctor's visit in the past year (36.5% versus 1974's 8.10%), but the disparity in engagement is still striking.

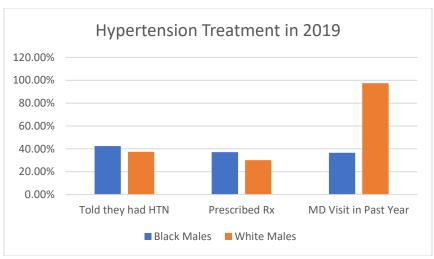


Image 1.3: Hypertension Treatment in 2019.

Interestingly, the communication regarding hypertension and the number of prescriptions written has increased in both populations over the past 45 years, and it seems evident from some of the changes made in the survey questions from 1974 to the present day that public health literacy surrounding conditions such as hypertension and cardiovascular disease were made a priority. The data also shows straightforward evidence that rates of diagnosed hypertension and, thus, the prescriptions to treat the condition are more prevalent today than they were in 1974 as well.

Assumptions and Limitations

A minor number of assumptions and adjustments had to be made in order to normalize the data from surveys taken in 1974 to surveys taken in 2019, explicitly surrounding the variables that required adjustments, such as adjusting income for inflation, the creation of a BMI calculator for the 1974 height and weight data, and data points surrounding whether the respondent had ever been told they had cardiovascular disease, which was simply called "heart trouble" in the 1974 survey. This speaks again to the health literacy challenge at the time.

Even though I organized and began the analysis across fourteen variables, the length, and depth of this project have limited the number of data points used in the findings shared. I do think this data set

has an incredible amount of potential, and given that it is a more extended project with more resources to hold and store data, it may highlight some progress in healthcare disparities as well as areas of opportunity for longer-term improvement.

The only data point used in the analyses above that required adjustment was the survey question regarding the timing of the last doctor's visit. I mapped the data to align with the 2019 survey values out of necessity, but we did lose a significant amount of granularity there in that the 1974 survey asked not only about doctor's visits in the past year, but also in the past six months, three months, four weeks, and two weeks. See image 1.4 and the associated screenshots below for the original verbiage.

1974 -	,	
127	Q19b	INTERVAL SINCE LAST DOCTOR VISIT O Never 1 Visit in past 2 weeks 3 2 weeks to less than 6 months 4 6 months to less than 12 months 5 1 year 6 2-4 years 7 5+ years 9 Unknown

2019 -

LASTDR_A Time since last saw doctor

Code	Description
0	Never
1	Within the past year (anytime less than 12 months ago)
2	Within the last 2 years (1 year but less than 2 years ago)
3	Within the last 3 years (2 years but less than 3 years ago)
4	Within the last 5 years (3 years but less than 5 years ago)
5	Within the last 10 years (5 years but less than 10 years ago)
6	10 years ago or more
7	Refused
8	Not Ascertained
9	Don't Know

Image 1.4: Original Mapping of Interval/Time Since Last Doctor Visit

The mapping used for the final dataset was:

- 0 Never
- 1 Within the past year (Includes 1974 responses of 1-4)
- 2 Within the past 2 -4 years (1974: 6 and 2019: 2-4)

- 3 Five years or greater (1974: 7 and 2019: 5-6)
- 4 Refused, Not Ascertained, or Unknown (1974: 9, 2019: 7-9)

Ethical Concerns

Ethical concerns are always present when studying population health. The surveys utilized here are intended to capture the whole nation's experiences and health markers. While there is a strong body of data here to analyze, we cannot assume that a comprehensive survey filtered and used to examine an underrepresented minority population would have the same results as directly engaging and resurveying the affected population members.

I would also like to mention that data sets for public health are often used to not only identify opportunities for improvement, like in the case of healthcare racial disparities but are sometimes used to help generate ideas for solutions. Again, this represents an ethical concern only if the affected population members are not directly engaged as critical stakeholders in the proposed solutions.

Finally, because the survey data is representative of live in-person, or telephone survey processes, there will always be some inherent bias on the part of the surveyors and those being surveyed. If a more extensive project were undertaken with this dataset, I would recommend investigating deeper into the processes and modalities used to obtain the survey responses.

Challenges/Opportunities

Finding data to examine that reveals some aspects of systemic discrimination in healthcare is not challenging to do. The challenge and opportunity presented by using data science in the healthcare industry is to use it to highlight areas where clear discriminatory patterns still exist. Presenting the information to healthcare clinicians and leaders in a way that helps them understand why the disparity still exists can be extremely challenging.

Healthcare data also has a large amount of variability, practice to practice. Much of healthcare data is dependent on the software systems that are in place to intake the data and clinicians often push

for free-text, individualized ways to tell the patient story, even as the administrators push for more discrete, reportable fields. There is a lot of data in the healthcare industry with a huge amount of potential, but the industry must invest in experts like data scientists, who can turn the data into actionable information for the organizations and, hopefully, one day, the entire country.

Finally, the biggest challenge was one that was highlighted by my work, even on this small project. The inclination to include too much data, to feel that so many individual data points may have a meaningful impact on the result, and by including too much, there may be noticeably clear messages lost in the noise. I spent a great deal of time analyzing, adjusting, and re-mapping data points that I never got the opportunity to use because my first filters carried the strongest and most straightforward message in alignment with my problem statement.

Conclusion

The healthcare industry has been aware of racial and ethnic disparities in healthcare outcomes for decades. Now, many social pressures are converging, and it may be possible to make some positive changes to this long-lasting problem. Communication of the data and how the related information will be presented is key.

The presence of a disparity, even one as specific as Black males with hypertension not getting appropriate levels of doctor's visits, does not have to be paired with a message that generates blame and defensiveness in the conference rooms of American hospitals and clinics. The data simply reveal patterns that show where treatment and access still have a lot of inequality across the system. This does not mean all our clinicians are to blame. The patterns revealed by the data show the impact of the barriers that are built into our processes and institutions in ways that can be invisible to those who are not living in the US as a minority.

If we can examine and measure the effects of these healthcare disparities, we have a greater chance of trying to intervene and change these processes and, ideally, then engage with the affected populations to brainstorm real solutions.

References:

- Tsao CW, Aday AW, Almarzooq ZI, et al. Heart disease and stroke statistics—2022 update: a report from the American Heart Association. Circulation. 2022;145(8):e153–e639.
 doi: 10.1161/CIR.0000000000001052.
- Aggarwal R, Chiu N, Wadhera RK, et al. Racial/ethnic disparities in hypertension prevalence, awareness, treatment, and control in the United States, 2013 to 2018. Hypertens Dallas Tex 1979. 2021;78(6):1719–1726. 10.1161/HYPERTENSIONAHA.121.17570.
- Hardy ST, Chen L, Cherrington AL, et al. Racial and ethnic differences in blood pressure among US adults, 1999–2018. Hypertension. 2021;78(6):1730–1741.
 doi: 10.1161/HYPERTENSIONAHA.121.18086.
- Bleich SN, Findling MG, Casey LS, Blendon RJ, Benson JM, SteelFisher GK, Sayde JM, Miller C.
 Discrimination in the United States: Experiences of Black Americans. Health Serv Res. 2019
 Dec;54 Suppl 2(Suppl 2):1399-1408. doi: 10.1111/1475-6773.13220. Epub 2019 Oct 29. PMID: 31663124; PMCID: PMC6864380.
- Assari S, Bazargan M. Minorities' Diminished Returns of Educational Attainment on Hospitalization Risk: National Health Interview Survey (NHIS). Hosp Pract Res. 2019
 Summer;4(3):86-91. doi: 10.15171/HPR.2019.17.
- 6. Stephens-Davidowitz, Seth (2017). Everybody Lies: Big Data, New Data, and What the Internet

 Can Tell Us about Who We Really Are. Dey St.

- 7. Shan, C., Wang, H., Chen, W. & Song, M (2015). *The Data Science Handbook*. The Data Science Bookshelf.
- 8. R. Hardeman, P. Homan, T. Chantarat, et al. *Improving the Measurement of Structural Racism to Achieve Antiracist Health Policy* (Feb 2022) Health Affairs.

 https://doi.org/10.1377/hlthaff.2021.01489
- United States Department of Health and Human Services. Centers for Disease Control and Prevention. National Center for Health Statistics. National Health Interview Survey,
 1974: Hypertension Supplement. Inter-university Consortium for Political and Social Research [distributor], 2010-11-29. https://doi.org/10.3886/ICPSR09796.v3

Appendix 1: 2019 and 1974 Variables Crosswalk

The following charts provide details about the values/responses used for each variable of both the 1974 and the 2019 datasets. Notes have also been added below any variable in which data was adjusted or normalized to allow for a single, analyzable dataset.

Column A: RACE

Year	Column Title	Tape Location/Code
1974	RACE:	50
2019	RACE:	RACEALLP_A_56

1974 -

	t	L
50	Q1	RACE (DETAIL)
		1 White 2 Negro 3 Other

2019 -

RACEALLP_A Single and multiple race groups

_		
Code	Description	
1	White only	
2	Black/African American only	
3	Asian only	
4	AIAN only	
5	AIAN and any other group	
6	Other single and multiple races	
7	Refused	
8	Not Ascertained	
9	Don't know	

Column B: SEX

Year	Column Title	Tape Location/Code
1974	SEX:	52
2019	SEX:	SEX_A_48

1974 –

52	Q1	SEX
		1 Male 2 Female

2019 –

SEX	Δ	Sex	of	Samp	P	Adult

Code	Description	
1	Male	
2	Female	
7	Refused	
8	Not Ascertained	
9	Don't Know	

Column C: EDUCATIONAL LEVEL

Year	Column Title	Tape Location/Code
1974	EDUCATIONAL LEVEL:	63
2019	EDUCATIONAL LEVEL:	EDUC_A_49

1974 -

EDUC_A Educational level of sample adult

Code	Description
00	Never attended/kindergarten only
01	Grade 1-11
02	12th grade, no diploma
03	GED or equivalent
04	High School Graduate
05	Some college, no degree
96	Associate degree: occupational, technical, or vocational program
07	Associate degree: academic program
08	Bachelor's degree (Example: BA, AB, BS, BBA)
09	Master's degree (Example: MA, MS, MEng, MEd, MBA)
10	Professional School degree (Example: MD, DDS, DVM, JD)
11	Doctoral degree (Example: PhD, EdD)
97	Refused
98	Not Ascertained
99	Don't Know

Final Mapping --

- 0 Not an HS Graduate or Equivalent
- 1 GED or HS Graduate
- 2 Some College
- 3 College Graduate of ANY level

Column D: Family Income

Year	Column Title	Tape Location/Code
1974	FAMILY INCOME RECODE:	70
2019	FAMILY INCOME RECODE:	FAMINCTC_A_616

1974 -

	1	1
70	Q38 Recode	FAMILY INCOME RECODE 1 Under \$3,000 2 \$3,000-4,999 3 5,000-6,999 4 7,000-9,999 5 10,000-14,999 6 15,000 + 7 Unknown -

2019 -

FAMINCTC A	Sample	adult	family	/ income
------------	--------	-------	--------	----------

Code	Description
000000-219999	\$0 to \$219,999
220000	\$220,000+

Family income in the 1974 dataset should be adjusted for inflation to combine the two datasets, and according to the US Bureau of Labor Statistics, the dollar had an average inflation rate of 3.79% per year between 1974 and 2019, producing a cumulative price increase of 518.88%. This means \$100.00 in 1974 is equivalent to \$618.88 of purchasing power in 2019.

This allowed me to create a formula in excel which adjusted the 1975 income to the 2019 standard.

Column E: Veteran Status

Year	Column Title	Tape Location/Code
1974	VETERAN STATUS:	67
2019	VETERAN STATUS:	FAMINCTC_A_616

1974 -

67	Q35	VETERAN STATUS
a mare		0 Female or under 17 years of age 1 Nonveteran 2 Peacetime only 3 World War I 4 World War II 5 Korean War 6 Viet Nam veteran 7 DK if served in Armed Forces
	I	8 DK if war veteran

2019 -

AFVET_A Ever serve active duty military

Code	Description	
1	Yes	
2	No	
7	Refused	
8	Not Ascertained	
9	Don't Know	

Veteran status is a field which required some adjustment to make the values comparable. Mapping used for this was:

- 1 Yes (Includes 2-6 in the 1974 responses)
- 2 No (Includes 0-1 in the 1974 responses)
- 9 Don't Know (Includes 7-8 in the 1974 responses)

Column F: Time Since Last Doctor Visit

Year	Column Title	Tape Location/Code
1974	TIME SINCE LAST DR VISIT	67
2019	TIME SINCE LAST DR VISIT	FAMINCTC_A_616

127	Q19b	INTERVAL SINCE LAST DOCTOR VISIT
		0 Never 1 Visit in past 2 weeks 3 2 weeks to less than 6 months 4 6 months to less than 12 months 5 1 year 6 2-4 years 7 5+ years 9 Unknown

2019 -

LASTDR_A Time since last saw doctor

Code	Description
0	Never
1	Within the past year (anytime less than 12 months ago)
2	Within the last 2 years (1 year but less than 2 years ago)
3	Within the last 3 years (2 years but less than 3 years ago)
4	Within the last 5 years (3 years but less than 5 years ago)
5	Within the last 10 years (5 years but less than 10 years ago)
6	10 years ago or more
7	Refused
8	Not Ascertained
9	Don't Know

Time Since Last Doctor Visit is a column that required some adjustment, as the 2019 survey captured longer segments of time than the 1974 survey.

Mapping used for the final dataset was:

- 0 Never
- 1 Within the past year (Includes 1974 responses of 1-4)
- 2 Within the past 2 -4 years (1974: 6 and 2019: 2-4)
- 3 Five years or greater (1974: 7 and 2019: 5-6)
- 4 Refused, Not Ascertained, or Unknown (1974: 9, 2019: 7-9)

Column G: Ever Told You Had Hypertension

Year	Column Title	Tape Location/Code
1974	EVER TOLD HTN:	466
2019	EVER TOLD HTN:	HYPEV_A_68

1974 –

466	1b.	EVER TOLD HAD HYPERTENSION
		1 Yes 2 No 3 DK or refused 4 Not applicable

2019 –

HYPEV_A	Ever been told you had hypertension		
Code	Description		
1	Yes		
2	Ne		

1	Yes	
2	No	
7	Refused	
8	Not Ascertained	
9	Don't Know	

This column did not have any adjustments made because an answer of "1 – Yes" is the only filter needed for the relevant queries.

Column H: Taking medication prescribed by a doctor for hypertension.

	· · · · · · · · · · · · · · · · · · ·	71
Year	Column Title	Tape Location/Code
1974	NOW TAKING RX:	477
2019	NOW TAKING RX:	HYPMED_A

1974 -	L	1
477	5Ъ	NOW TAKING MEDICINE 1 Yes 2 No 3 Unknown 4 Not advised to take medicine bl Not applicable

2019 –

HYPMED	Δ	Now	taking	high	hlood	nressure	medication
RIPPED	A	NOW	CONTINE	HITEH	DIOOG	pressure	mentcatton

Code	Description	
1	Yes	
2	No	
7	Refused	
8	Not Ascertained	
9	Don't Know	

Columns I, J, and K: Weight, Height, and BMI

Year	Column Title Tape Location/Co	
1974	WEIGHT (lbs.)	513
1974	HEIGHT (inches):	515
1974	CALCULATED BMI:	N/A
2019	HEIGHT (lbs.)	HEIGHTTC_A_198
2019	WEIGHT (lbs.)	WEIGHTLBTC_A_200
2019	BMI	BMICAT_A_203

1974 Weight & Height –

		l
515-517	17a	WEIGHT 000-996 0-996 pounds 997 997+ pounds 998 Unknown
518-519	17b	HEIGHT 12-97 inches 98 98+ inches 99 Unknown

2019 – Height, Weight, and BMI

HEIGHTTC_A Total height in inches, public use

Code	Description	
59-76	59-76 inches	
96	Not available	
97	Refused	
98	Not Ascertained	
99	Don't Know	

WEIGHTLBTC_A Weight without shoes (pounds), public use

Code	Description	
100-299	100-299 pounds	
996	ot available	
997	Refused	
998	Not Ascertained	
999 Don't Know		

BMICAT_A	Categorical	body	mass	index,	public	use
----------	-------------	------	------	--------	--------	-----

Code	Description	
1	nderweight	
2	althy weight	
3	Overweight	
4	Obese	
9	Unknown	

The only adjustment made was due to the lack of BMI calculation in the 1974 dataset. An additional column added included a formula that calculated the BMI based on the height and weight records present: =703 * WEIGHT/HEIGHT ^ 2.

The 2019 dataset gave a category for the BMI calculated, but did not include the exact BMI result, so the formula was applied to the 2019 data in the same way.

Column L: Been Told You Have Diabetes?

1974	DIABETES:	525
2019	DIABETES:	DIBEV_A_181



		L
525	18	DIABETES
		1 Yes 2 No 3 Unknown
	1	,

2019 –

DIBEV_A	Ever	had	diabetes
---------	------	-----	----------

Code	Description	
1	Yes	
2	No	
7	Refused	
8	Not Ascertained	
9	Don't Know	

Column M: Been Told You Had "Heart Trouble" or "Coronary Artery Disease"

Year	Column Title	Tape Location/Code
1974	DIABETES:	526
2019	DIABETES:	CHDEV_A

1974 –

526	19	HEART TROUBLE
		1 Yes 2 No 3 Unknown

2019 –

CHDEV_A Ever been told you had coronary heart disease

Code	Description
1	Yes
2	No
7	Refused
8	Not Ascertained
9	Don't Know

Column N: Had A Stroke

Year	Column Title	Tape Location/Code
1974	STROKE:	527
2019	STROKE:	STREV_A_78

1974 – 527 20 STROKE 1 ... Yes 2 ... No

2019 -

STREV_A	Ever been told you had a stroke
Code	Description
1	Yes
2	No
7	Refused
8	Not Ascertained
9	Don't Know

Column Removed: AGE

Year	Column Title	Tape Location/Code
1974	N/A:	94
2019	N/A:	AGEP_A_45

3 ... Unknown

2019 -

AGEP_A	Age	of	SA	(top	coded)
- 1					

Code	Description	
18-84	18-84 years	
85	85+ years	
97	Refused	
98	Not Ascertained	
99	Don't Know	

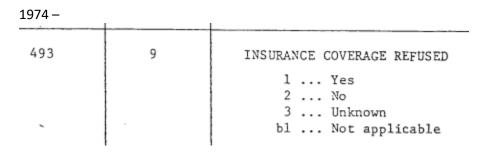
1974 -

137 .				
94	R	AGE OF RESPONDENT		
		1 Under 20 years 2 20-54 years 3 55-64 years 4 65-74 years 5 75+ years 6 Unknown and not reported		

The age column was removed from the dataset because the age categories created in the 2019 survey reduced population segmentation to the point that filtering this data would not have meaning to my query.

Column Removed: Insurance Coverage Refused to Cover Treatment

Year	Column Title	Tape Location/Code
1974	N/A:	493
2019	N/A:	AGEP_A_45



Although this question was removed from both datasets, it does highlight the increased complexity of the healthcare delivery system over the past 45 years. In 1974, one question was asked, "Did your insurance refuse to cover treatment of your hypertension?". In the 2019 survey, there were 124 individual questions regarding insurance payor, coverage, deductibles, and other details. There was no single question that would have mapped easily to the 1974 query.