SPECIAL FOCUS ISSUE: BLOOD PRESSURE

Generalizing Intensive Blood Pressure Treatment to Adults With Diabetes Mellitus

Seth A. Berkowitz, MD, MPH,[a,b] Jeremy B. Sussman, MD, MS,[c,d] Daniel E. Jonas, MD, MPH,[a,b] Sanjay Basu, MD, PhD[e,f]

ABSTRACT

BACKGROUND Controversy over blood pressure (BP) treatment targets for individuals with diabetes is in part due to conflicting perspectives about generalizability of available trial data.

OBJECTIVE The authors sought to estimate how results from the largest clinical trial of intensive BP treatment among adults with diabetes would generalize to the U.S. population.

METHODS The authors used transportability methods to reweight individual patient data from the ACCORD (Action to Control Cardiovascular Risk in Diabetes) BP trial (N = 4,507) of intensive (goal systolic BP <120 mm Hg) versus standard (goal systolic BP <140 mm Hg) treatment to better represent the demographic and clinical risk factors of the U.S. population of adults with diabetes (data from NHANES [National Health and Nutrition Examination Survey] 2005 to 2014, n = 1,943). The primary outcome was the first occurrence of nonfatal myocardial infarction, nonfatal stroke, or cardiovascular death. Analysis used weighted Cox proportional hazards regression models with robust standard errors.

RESULTS The ACCORD BP sample had less racial/ethnic diversity and more elevated cardiovascular risk factors than the NHANES participants. Weighted results significantly favored intensive BP treatment, unlike unweighted results (hazard ratio for primary outcome in intensive versus standard treatment in weighted analyses: 0.67, 95% confidence interval: 0.49 to 0.91; in unweighted analyses: hazard ratio: 0.88, 95% confidence interval: 0.73 to 1.07). Over 5 years, the weighted results estimate a number needed to treat of 34, and number needed to harm of 55.

CONCLUSIONS After reweighting to better reflect the U.S. adult population with diabetes, intensive BP therapy was associated with significantly lower risk for cardiovascular events. However, data were limited among racial/ethnic minorities and those with lower cardiovascular risk. (J Am Coll Cardiol 2018;72:1214–23) © 2018 by the American College of Cardiology Foundation.

Listen to this manuscript's audio summary by *JACC* **Editor-in-Chief Dr. Valentin Fuster.**

https://doi.org/10.1016/j.jacc.2018.07.012

Uncontrolled blood pressure (BP) remains a leading cause of excess cardiovascular morbidity and mortality among patients with diabetes mellitus (DM) (1). Determining the appropriate BP control target is of major clinical and public health importance. However, the BP targets to adopt for this population remain controversial, with conflicting recommendations from different guideline-issuing groups. The 2017 update of the American College of Cardiology (ACC)/American Heart Association (AHA) BP treatment guidelines (2) recommend universal intensive BP treatment for adults with DM (target BP <130/80 mm Hg). The 2018 Standards of Care in Diabetes from the American Diabetes Association, however, recommend a target of <140/90 mm Hg for most patients (3). Similarly, the American College of Physicians and the American Academy of Family Physicians have not endorsed the ACC/AHA hypertension guidelines for individuals with DM (4). The differences among guidelines leave individual practitioners and patients with a dilemma when deciding on a target BP.

The differences among current guidelines largely result from different interpretations of how to generalize 2 large randomized trials. The SPRINT trial (Systolic Blood Pressure Intervention Trial) (5) enrolled high-risk individuals without DM to intensive versus standard BP control (systolic target <120 mm Hg vs. <140 mm Hg) and found a 27% reduction in mortality for intensive BP treatment. Yet the largest randomized clinical trial of intensive BP treatment among people *with* DM, the ACCORD BP (Action to Control Cardiovascular Risk in Diabetes–Blood Pressure) trial did not find significant benefit of intensive BP treatment for its primary outcome (a composite of nonfatal myocardial infarction (MI), nonfatal stroke, or death from cardiovascular causes; hazard ratio [HR]: 0.88, 95% confidence interval [CI]: 0.7 to 1.06) (6). Further, participants in the intensive BP treatment arm of the ACCORD BP trial experienced more serious adverse events related to BP medications than participants in the standard arm (serious adverse event rate 3.3% vs. 1.3%; p < 0.0001) (6).

A new possibility to inform clinical decisions in the presence of conflicting trial results is to quantitatively assess the generalizability of a trial to the population of patients among whom the clinical decision is meant to apply. Recent advances in statistical approaches called "transportability methods" allow us to quantify how the results of the ACCORD BP trial may generalize to the broader U.S. population of patients with DM, and specifically estimate how

much the general population might experience different outcomes from intensive BP treatment than the result seen in the ACCORD BP trial (7-11). Transportability methods can be thought of as reweighting the results from each member of the trial sample to construct a pseudopopulation that matches the demographics and clinical characteristics of the broader national population with DM. By weighting the trial results, the overall effectiveness of intensive BP treatment for the more general population can be estimated, helping to inform whether individual practitioners may be wary or eager to implement intensive BP treatment among patients with DM.

Here, we sought to identify how the ACCORD BP study results may be generalized to people with DM in the United States, and thereby inform recommendations for BP treatment.

METHODS

SOURCE OF DATA AND SAMPLE SIZE. Individual patient data from the ACCORD BP trial (n = 4,507) (6), and individual-level data pooled from repeated cross sections of the NHANES (National Health and Nutrition Examination Survey, 2005 to 2014, n = 1,943), a nationally representative population-based epidemiological surveillance study of noninstitutionalized Americans (12) were used to conduct the study. Analyses incorporated NHANES design information, including sampling weights and clustering, as appropriate. We did not use data from the SPRINT trial in this study because the SPRINT trial did not include individuals with DM, and thus data on several relevant cardiovascular risk factors for individuals with diabetes were not measured. Further, the SPRINT trial did not release socioeconomic status information that would be used for transportability estimates (5).

PARTICIPANTS. ACCORD BP trial. Detailed information on the study design and main outcomes of the ACCORD BP trial were previously published (6). In brief, the ACCORD trial enrolled 10,251 type 2 diabetes patients with a history of a hemoglobin A_{1c} (HbA_{1c}) >7.5%, and cardiovascular disease (if age >40 years) or high cardiovascular risk (if age >55 years) to an intensive versus standard glycemic control strategy. Using a 2 × 2 factorial design, patients were additionally assigned to either a lipid-lowering study (the ACCORD lipid trial) or a BP-lowering study (ACCORD BP, N = 4,733). Additional entry criteria for the ACCORD BP trial are described in the study protocol,

TABLE 1 Demographic and Clinical Characteristics of NHANES and ACCORD Participants

	ACCORD Before Weighting (n = 4,507)	p Value	ASD	NHANES (n = 1,943, Representing 26,733,959)	ACCORD After Weighting (n = 4,507)	p Value	ASD
Baseline age, yrs	62.76 ± 6.70	<0.001	0.30	59.65 ± 13.70	60.87 ± 6.59	0.002	0.13
Female	47.4	0.42	0.03	48.8	48.4	0.87	0.01
Race/ethnicity		<0.001	0.34			<0.001	0.34
Non-Hispanic white	58.9			61.8	45.5		
Non-Hispanic black	23.9			15.3	19.6		
Hispanic	6.9			15.4	25.5		
Asian/multi/other	10.4			7.5	9.4		
Education		<0.001	0.29			<0.001	0.20
Less than HS	16.2			26.8	31.9		
HS diploma	26.8			27.0	19.1		
Some college	32.5			29.1	29.1		
College diploma or higher	24.5			17.1	19.9		
Insured	84.6	0.91	0.07	86.8	84.9	0.18	0.06
Smoking status		<0.001	0.18			0.03	0.11
Never	44.9			51.1	56.7		
Current	13.1			15.4	13.6		
Former	42.0			33.4	29.7		
History of CHF	4.3	<0.001	0.19	9.0	11.1	0.26	0.07
History of MI	13.8	0.002	0.10	10.5	11.0	0.68	0.02
History of stroke	6.4	0.02	0.08	8.4	7.9	0.70	0.02
Years with diabetes	10.86 ± 7.84	<0.001	0.38	7.49 ± 9.19	10.53 ± 7.56	<0.001	0.35
BMI, kg/m^2	32.16 ± 5.46	0.003	0.11	32.80 ± 7.30	32.62 ± 5.87	0.37	0.04
SBP, mm Hg	139.23 ± 15.83	<0.001	0.51	130.08 ± 19.15	132.93 ± 15.04	<0.001	0.16
DBP, mm Hg	76.01 ± 10.35	<0.001	0.62	69.51 ± 12.96	71.88 ± 10.05	<0.001	0.31
FPG, mg/dl	174.90 ± 57.82	<0.001	0.33	151.87 ± 54.57	180.23 ± 66.36	<0.001	0.39
HbA$_{1c}$, %	8.34 ± 1.09	<0.001	0.78	7.16 ± 1.64	8.07 ± 1.03	<0.001	0.59
HDL, mg/dl	46.16 ± 13.64	<0.001	0.15	49.11 ± 13.45	50.87 ± 16.36	<0.001	0.17
LDL, mg/dl	109.93 ± 36.75	<0.001	0.16	103.86 ± 36.02	105.44 ± 34.10	0.41	0.04
Triglycerides, mg/dl	193.70 ± 176.29	0.001	0.11	148.89 ± 76.09	127.11 ± 68.86	<0.001	0.35
Estimated GFR, ml/min	91.54 ± 28.93	<0.001	0.15	87.46 ± 28.09	84.12 ± 21.68	0.003	0.12
Urine albumin to creatinine ratio	93.15 ± 326.11	0.63	0.02	75.41 ± 481.22	197.62 ± 739.92	0.004	0.17

Values are mean ± SD or %. Means and percentages for NHANES are nationally representative using NHANES sampling weights and clustering variables. p and ASD values are in comparison to NHANES.

ASD = absolute standardized difference; BMI = body mass index; CHF = congestive heart failure; DBP = diastolic blood pressure; FPG = fasting plasma glucose; HbA$_{1c}$ = hemoglobin a$_{1c}$; HDL = high-density lipoprotein cholesterol; HS = high school; GRF = glomerular filtration rate, estimated by the modification of diet in renal disease equation; LDL = low-density lipoprotein cholesterol; MI = history of myocardial infarction; NHANES = National Health and Nutrition Examination Survey; SBP = systolic blood pressure.

but include having a systolic blood pressure (SBP) between 130 and 180 mm Hg on 3 or fewer BP-lowering medications, and urinary protein excretion <1.0 g in 24 h. Participants in the ACCORD BP trial were randomly assigned to the intensive BP treatment (n = 2,362; goal SBP <120 mm Hg; achieved mean SBP after 1 year 119.3 mm Hg) or standard BP treatment (n = 2,371; goal SBP <140 mm Hg; achieved mean SBP after 1 year 133.5 mm Hg) arm. Treatment was not blinded, and strategy for achievement of that the treatment goal (e.g., choice of medications and dose adjustment) was determined by individual study physicians under broad guidance. The maximum duration of follow-up was 7 years.

NHANES. To determine the sociodemographic and clinical characteristics of the civilian U.S. population of patients with DM, we used demographic, questionnaire, examination, and fasting laboratory data from NHANES pooled across 10 years (2005 to 2014), which incorporates the most recent data available. As in prior studies, DM was defined as patient self-report of a diagnosis of diabetes, HbA$_{1c}$ >6.5%, or fasting plasma glucose (FPG) >126 mg/dl (13-15). We did not distinguish between type 1 and type 2 diabetes in this study because current BP guidelines do not. Because current guidelines are meant to apply to adults, we included NHANES participants ≥20 years of age, and because separate guidelines exist for management of BP in pregnancy, we excluded those currently pregnant. Because we were concerned that data within NHANES may not distinguish between those with controlled hypertension

CENTRAL ILLUSTRATION Intensive Blood Pressure Control in Diabetes
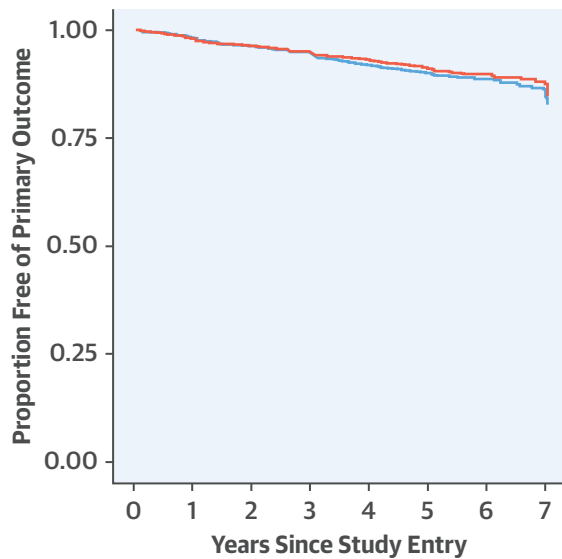
The Kaplan-Meier estimates of experiencing the ACCORD BP primary outcome (a composite of first occurrence of nonfatal myocardial infarction, nonfatal stroke, or cardiovascular death), in the unweighted **(left)** and weighted **(right)** analyses. ACCORD = Action to Control Cardiovascular Risk in Diabetes trial; BP = blood pressure.
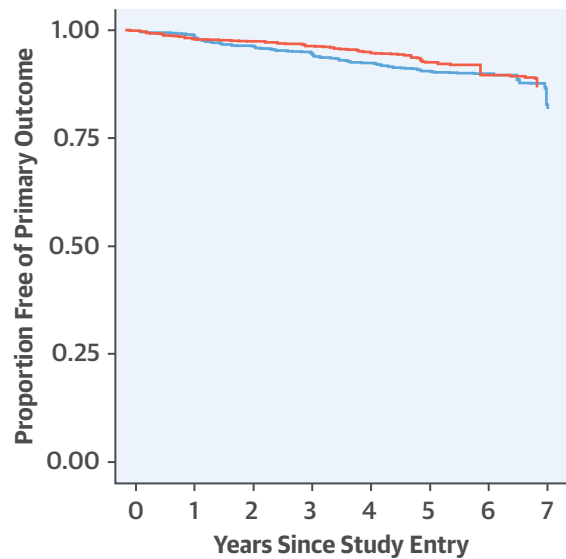
Berkowitz, S.A. et al. J Am Coll Cardiol. 2018;72(11):1214–23.

from those without hypertension, we did not have explicit hypertension-related inclusion criteria in our main analyses. However, to investigate whether this decision could affect the results, we performed 2 sensitivity analyses. In the first, in addition to the aforementioned criteria, we also required a self-report of hypertension diagnosis, a self-report of taking a BP-lowering medication, a SBP measurement (averaged over all NHANES measurements) of >130 mm Hg, or a diastolic BP measurement (averaged over all NHANES measurements) of >80 mm Hg (2). The second sensitivity analysis sample required that SBP (averaged over all NHANES measurements) be between 130 and

180 mm Hg, which closely matches an ACCORD BP eligibility criterion (6).

OUTCOMES. The outcomes for this study matched those of the ACCORD BP trial. The primary outcome was a composite of first occurrence of nonfatal MI, nonfatal stroke, or cardiovascular death. Secondary outcomes used in this study were all-cause mortality, total (fatal and nonfatal) stroke, the main microvascular composite outcome from the ACCORD trial (first occurrence of renal failure [initiation of dialysis or end-stage renal disease, renal transplantation, or rise of serum creatinine >291.72 μmol/l] or retinal photocoagulation or vitrectomy to treat retinopathy), and number of serious nonhypoglycemic adverse events (defined as hypotension, syncope, bradycardia or arrhythmia, hyperkalemia, or renal failure, attributable to the study medication) (6).

FACTORS USED FOR WEIGHTING. We considered several potential effect modifiers that may alter the risk or benefit of intensive versus standard BP treatment. These factors were selected on the basis of prior evidence or hypotheses for their association with cardiovascular disease risk (3,16–20). The factors considered were age, sex, race/ethnicity (categorized as non-Hispanic white, non-Hispanic black, Hispanic, and other/multiracial), education (categorized as less than high school diploma, high school diploma, some college, or college degree and higher), health insurance status (insured/uninsured), tobacco smoking status (never, current, former), history of MI, history of congestive heart failure, history of stroke, SBP, diastolic BP, total cholesterol, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, triglycerides, HbA$_{1c}$, FPG, glomerular filtration rate (estimated using the 4-term modification of diet in renal disease equation), urinary albumin to creatinine ratio, and body mass index (Table 1). All potential effect modifier measurements from the ACCORD BP study were taken from baseline (pre-treatment) examinations. Data from the ACCORD BP trial and NHANES were inspected for implausible values, but only a small number of diastolic BP readings <30 mm Hg from NHANES were identified and excluded from the analysis.

MISSING DATA. Missing data were not imputed because missingness was <5% for any variable.

STATISTICAL ANALYSIS. To generalize the ACCORD BP results to the NHANES population (i.e., to "transport" the results from the ACCORD BP trial to NHANES), we used a stabilized inverse odds of selection weighting approach that is analogous to propensity score weighting (8). The intuition behind

TABLE 2 Relative Risk of Outcomes for Intensive Versus Standard BP Therapy

	Unweighted		Weighted	
	HR/IRR	95% CI	HR/IRR	95% CI
Primary outcome	0.88	0.73-1.07	0.67	0.49-0.91
Total mortality	1.07	0.85-1.35	0.91	0.62-1.32
Stroke	0.59	0.39-0.89	0.33	0.17-0.62
Microvascular outcome	1.07	0.90-1.27	1.00	0.75-1.33
Serious adverse events	1.93	1.37-2.72	1.97	1.09-3.58

All estimates represent hazard ratios (HRs) and 95% confidence intervals (CIs) from proportional hazards regression models except for serious adverse events which represents incidence rate ratios from negative binomial models. Models are adjusted for intensive versus standard glycemic therapy, clinical network, and history of cardiovascular disease at baseline.

BP = blood pressure; IRR = incidence rate ratio.

this approach is to up-weight individuals in the ACCORD BP trial who have characteristics that are more common in NHANES, and down-weight those in the ACCORD BP trial with characteristics less common in NHANES (Central Illustration). To calculate the weights, the probability of not being included in the ACCORD trial was calculated for each person in NHANES, conditional on the aforementioned covariates, divided by their individual probability of being in the ACCORD trial, conditional on the aforementioned covariates. The odds were then multiplied by a stabilization factor, which was the unconditional probability of being included in the ACCORD trial divided by the unconditional probability of not being included. In mathematical terms, the stabilized inverse odds of selection were:

$$(1 - p|Z)/(p|Z) \cdot (p/1 - p)$$

where p is the probability of selection and Z is the vector of individual covariates. To estimate the conditional probability of selection, we fit a logistic regression model (Online Table 1) with the outcome of being in the ACCORD trial and the above covariates; NHANES sampling weights were then multiplied by the calculated selection weights to generate a nationally-representative population. The unconditional probability of selection was based on a null (intercept-only) logistic regression model, again incorporating NHANES weights. As a robustness check, we re-estimated the probabilities using a machine learning approach that can capture nonlinearity between covariates and the outcome and complex interactions among covariates in a more sophisticated manner than standard logistic regression (Online Table 2) (21). Weights were Winsorized at the 2.5 and 97.5 percentiles to reduce the impact of extreme weights (22,23).

Outcome metrics. We calculated outcome metrics on both relative and absolute scales. For results on a relative scale, we used the same analytic approach as the main the ACCORD BP trial analysis, fitting proportional hazard (Cox) regression models that contain terms for intensive versus standard BP treatment, adjusting for intensive versus standard glycemic treatment arm, clinical center, and previous history of cardiovascular event (6). A previous study revealed that the proportional hazards assumption was reasonable in the ACCORD BP trial (6). We present unweighted results, which replicate the results of the main ACCORD BP paper as a check on our analysis, and results weighted by the inverse odds of selection to produce estimates transported to the NHANES population. Because the version of the ACCORD trial data available to us provides counts of serious adverse events rather than exact time-to-event data (to protect confidentiality), we used unweighted and weighted negative binomial models, offset by the person-time of follow-up for all-cause mortality, to model serious adverse events. Robust standard errors with independent covariance structure were estimated to calculate 95% CIs.

For results on the absolute scale, we calculated number of events per 1,000 person-years of follow-up, and plotted the survival function for time-to-event outcomes using Kaplan-Meier methods. We applied log-rank tests for significance of time-to-event outcomes, and the Wilcoxon 2-sample test (Mann-Whitney *U* test) for serious adverse events with a design correction factor for weighted analyses.

Sensitivity and robustness checks. First, we conducted Cox/negative binomial regression analyses adjusting for factors that had an absolute standardized difference >0.2 after weighting, to correct for residual imbalance between the weighted ACCORD BP trial population and the NHANES population. Second, we repeated analyses using a targeted maximum likelihood estimation approach, which has been shown to be less sensitive to positivity violations (cases where some combinations of important covariates are not observed in both the ACCORD BP and NHANES datasets) (9,10). Next, we fit nested conditional probability models to identify combinations of characteristics that might be responsible for differences in transported and untransported treatment effects. These models sequentially added subsets of the full set of cardiovascular risk factors based on imbalance between the ACCORD BP sample and the overall population. Finally, we fit the same weighted proportional hazards models in samples that additionally required hypertension-related inclusion criteria, as described above, to be met.

TABLE 3	Absolute Differences in Outcomes for Intensive Versus Standard BP Treatment					
	Unweighted			**Weighted**		
	Events per 1,000 Person-Years	**Difference**	**p Value**	**Events per 1,000 Person-Years**	**Difference**	**p Value**
Primary outcome						
Standard	21.27	−2.29	0.24	20.18	−5.77	0.03
Intensive	18.98			14.40		
Total mortality						
Standard	12.15	0.83	0.50	11.82	−0.80	0.72
Intensive	12.99			11.02		
Stroke						
Standard	5.41	−2.20	0.01	5.42	−3.68	0.002
Intensive	3.21			1.74		
Microvascular						
Standard	24.50	2.00	0.41	26.80	−0.01	0.98
Intensive	26.50			26.79		
Serious adverse events						
Standard	4.47	4.18	0.0003	3.41	3.64	0.01
Intensive	8.66			7.04		

p values from log-rank test, except for serious adverse events, which used Wilcoxon. Weighted analyses used design-corrected versions.
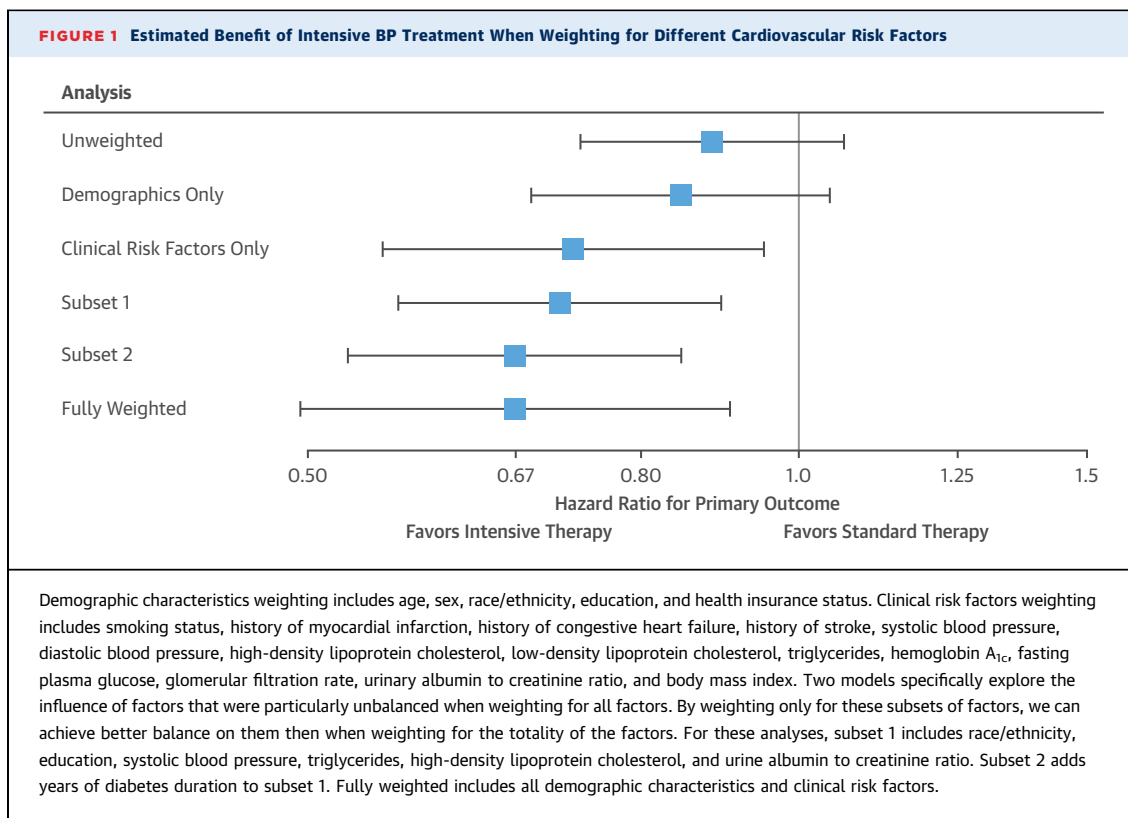
BP = blood pressure.

All analyses followed the intention-to-treat principle for treatment group assignment. A 2-sided p value <0.05 was taken to indicate statistical significance. Analyses were conducted in SAS 9.4 (SAS Institute, Cary, North Carolina) (PROC LOGISTIC and PROC PHREG) and R 3.4.2 (R Foundation for Statistical Computing, Vienna, Austria). Statistical code for replication is available from the authors.

The human research committee at Partners Healthcare approved this study of secondary data.

RESULTS

STUDY PARTICIPANT CHARACTERISTICS. Overall, 4,507 participants from the ACCORD BP study were included, consisting of 2,261 in the intensive treatment arm and 2,246 in the standard arm. The median follow-up time for the primary outcome was 4.92 years. There were 1,943 adults with diabetes included from the NHANES dataset. Comparisons between ACCORD BP and NHANES participants revealed significant differences between the study groups. In particular, ACCORD BP participants had less racial/ethnic diversity, greater education, and more elevated cardiovascular risk factors than the overall population (Table 1).

TRANSPORTED RESULTS FROM ACCORD BP TO NHANES. After weighting the ACCORD BP population to NHANES, the weighted ACCORD BP sample was more representative of the overall population, but as would be expected by the study design, some

FIGURE 1 Estimated Benefit of Intensive BP Treatment When Weighting for Different Cardiovascular Risk Factors

Demographic characteristics weighting includes age, sex, race/ethnicity, education, and health insurance status. Clinical risk factors weighting includes smoking status, history of myocardial infarction, history of congestive heart failure, history of stroke, systolic blood pressure, diastolic blood pressure, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, triglycerides, hemoglobin A_{1c}, fasting plasma glucose, glomerular filtration rate, urinary albumin to creatinine ratio, and body mass index. Two models specifically explore the influence of factors that were particularly unbalanced when weighting for all factors. By weighting only for these subsets of factors, we can achieve better balance on them then when weighting for the totality of the factors. For these analyses, subset 1 includes race/ethnicity, education, systolic blood pressure, triglycerides, high-density lipoprotein cholesterol, and urine albumin to creatinine ratio. Subset 2 adds years of diabetes duration to subset 1. Fully weighted includes all demographic characteristics and clinical risk factors.

imbalances remained, particularly in terms of racial/ethnic diversity, HbA$_{1c}$, FPG, and duration of diabetes (Table 1). In general, weighting increased the proportion of individuals who were Hispanic, and those who had lower HbA$_{1c}$, fasting glucose, and shorter duration of diabetes along with other changes as indicated in Table 1. The weights ranged from 0.01 to 4.67 (Online Figure 1).

In analyses examining the primary outcome, weighted results favored intensive BP treatment, unlike the unweighted results (HR for the primary outcome in intensive versus standard treatment in weighted analyses 0.67; 95% CI: 0.49 to 0.91) (Table 2). In absolute terms, the transported results would indicate 14.4 cardiovascular events in the intensive BP treatment arm versus 20.2 in the standard arm per 1,000 person-years (p = 0.03), with a number needed to treat of 172 over 1 year and 34 over 5 years to avert 1 cardiovascular event (Table 3). These contrast with the unweighted results of 21.27 events per 1,000 person-years in the standard arm and 18.98 in the intensive arm (p = 0.24).

In the weighted analyses, total mortality (HR: 0.91; 95% CI: 0.62 to 1.32) was not different between intensive and standard therapy groups, nor was the risk of microvascular events (HR: 1.00; 95% CI: 0.75 to 1.33). Stroke outcomes favored intensive therapy (HR:

0.33; 95% CI: 0.17 to 0.62), showing more favorable point estimates in weighted versus unweighted analyses (unweighted HR: 0.59; 95% CI: 0.39 to 0.89). Weighted and unweighted Kaplan-Meier plots showing absolute outcome metrics for the time-to-event outcomes are presented in Online Figures 2 to 9 and were consistent with the relative outcome metrics. Serious adverse events were more common in the weighted intensive therapy group (incidence rate ratio 1.97; 95% CI: 1.09 to 3.58, number needed to harm of 275 at 1 year or 55 at 5 years), and the estimate was similar to that of the unweighted analyses (unweighted incidence rate ratio 1.93; 95% CI: 1.37 to 2.72).

In sensitivity analyses that adjusted for residually imbalanced factors (race/ethnicity, education, diastolic BP, FPG, HbA$_{1c}$, triglycerides, and years with diabetes), point estimates were more strongly in favor of intensive BP therapy than in the weighted but unadjusted analyses (Online Table 3), though adjustment did not produce any qualitative changes. In sensitivity analyses additionally requiring hypertension diagnosis or elevated BP as inclusion criteria, estimates were similar to the main analyses (Online Tables 4 and 5). In sensitivity analyses using targeted maximum likelihood estimation, estimates were again more strongly in favor of intensive BP

therapy (Online Table 6). In analyses using nested models, weighting for demographic factors or clinical risk factors only did not fully explain the differences between unweighted and weighted results, but the combination of demographic and clinical factors (particularly race/ethnicity, education, SBP, triglycerides, high-density lipoprotein cholesterol, urine albumin to creatinine ratio, and years of diabetes duration) yielded results most similar to the "fully" weighted results (**Figure 1,** Online Tables 7 and 8).

DISCUSSION

We sought to transport the results of the ACCORD BP trial to a population more representative of Americans with diabetes, and found that intensive BP therapy would be expected to be associated with lower risk for the primary outcome of nonfatal MI, nonfatal stroke, or cardiovascular death, along with lower risk for stroke alone, in the general DM population. The estimated benefit of intensive treatment seen after reweighting is in line with those observed in the SPRINT trial, which helps harmonize the results of the 2 studies (5). These findings favor intensive BP treatment, but also highlight the lack of data support for BP treatment guidance in large segments of the U.S. adult population with DM—particularly racial/ethnic minorities and those with lower cardiovascular risk.

The magnitude of the estimated reduction in the primary outcome suggested a number needed to treat of 172 at 1 year and 34 at 5 years to avert one cardiovascular event. For reference, the ACC/AHA guidelines recommend treatment in those with a 10-year risk >10%, which will have a number needed to treat of around 300 at 1 year (2,24). Intensive therapy was also associated with reduced stroke incidence in weighted estimates, similar to unweighted estimates, but no difference in all-cause mortality or microvascular complications of diabetes. Intensive therapy was associated with a number needed to harm of 275 at 1 year.

These findings offer significant new contributions to the published reports on BP guidelines, and to the published reports on the interpretation of clinical trial results more broadly. Current guidelines for BP treatment among patients with diabetes do not restrict their recommendations to those included in major trials, and vary notably in their target BP recommendations (2). Guideline committees and individual practitioners do not have the luxury of waiting for a randomized controlled trial that perfectly matches the general U.S. population or a particular patient panel before making treatment recommendations or decisions; because randomized controlled trials are expensive and time consuming, and it is sometimes difficult to enroll certain groups, attempting to generalize trial results among a selected participant population to a more real-world population is inevitable. This is not a criticism; in the absence of a directly applicable trial, the alternative to transporting results is not to use trial results at all. However, recent statistical advances have allowed us to transport results quantitatively and formally (7-9), rather than in an informal or qualitative way, which may help advise clinical practice guidelines and help inform practitioners faced with diverging guidelines.

Our work also informs the specific interpretation of the SPRINT and ACCORD BP trials. Prior work analyzing who benefits most from intensive BP treatment found that non-Hispanic black race/ethnicity was associated with greater benefit (25). Another reanalysis of ACCORD BP data found that, among a subset of ACCORD BP participants who would have been eligible for SPRINT (apart from DM), intensive BP treatment reduced a composite outcome of cardiovascular events (26). Finally, prior work found that the relative benefit of treatment may be greater in those with lower, compared with higher, cardiovascular risk (27). These results are all consistent with the current study, where we observed more relative benefits of intensive BP treatment when the overall risk was lower. Conversely, this finding does differ from a meta-analysis of BP trials that did not find different relative risks associated with BP treatment as cardiovascular risk varied (24). However, that meta-analysis did not include ACCORD BP or SPRINT data. Further, subgroup analyses of the ACCORD BP trial in the original paper (6) suggested larger benefits with intensive therapy in individuals with lower baseline HbA$_{1c}$ and lower diastolic BP (though interaction terms were not significant). Because the U.S. population with diabetes has lower HbA$_{1c}$ and BP than the included sample, reweighting the study likely emphasized the experience of those participants.

STUDY LIMITATIONS. Transport methods can only standardize results over variables that have been measured. If data on important risk factors, such as social factors that may modify the effectiveness of therapy in the real world, are not available, and differ between the original and transported population, then the results may not fully reflect what would happen if the intervention was applied to the new population. Design choices about who to include in the ACCORD sample also make it more difficult to apply results to the overall U.S. population, as indicated by residual lack of balance for some factors.

Because the ACCORD BP trial specifically recruited individuals at high cardiovascular risk (higher than average for an individual with diabetes), there is poor overlap between the ACCORD BP trial and the U.S. population for some low risk groups. This leads to uncertainty in estimating the potential effects of intensive BP control, and should prompt caution in interpreting the study results until they can be confirmed in samples that better match the overall U.S. population of adults with DM. Further, the publicly available data on serious adverse events does not make it possible to attribute specific adverse events to BP (as opposed to glycemia) medications, or to conduct time-to-event analyses for this outcome.

The results of this study, in light of the new ACC/AHA guidelines (2), bring up the unanswered question of what to do when transported results qualitatively differ from those of the original trial. Given the novelty of transportability methods, we believe it is best to view the results of this study as hypothesis generating until more rigorous prospective validation of transportability methods, such as post-transportability clinical studies, enable us to formally assess outcomes among patients treated with and without clinical decisions informed by transportability analyses. Next, the residual imbalance between the weighted ACCORD BP and NHANES populations does highlight areas where the evidence favoring intensive BP has less data support. To address this, we conducted sensitivity analyses adjusted for residual imbalance between the weighted ACCORD BP and NHANES populations, and we used methods that were less sensitive to the unrepresentativeness of the ACCORD BP trial. In both cases, results were even more in favor of intensive BP therapy than our base case analysis; this suggests that residual confounding and lack of representativeness of the ACCORD BP trial are unlikely to explain our observed results. However, future studies should examine further the robustness of the intensive BP effect among racial/ethnic minorities with lower baseline cardiovascular risk, who were the least-informed by the ACCORD BP results. This brings up an important gap in knowledge for future research:

how much to transport results when data that directly answer the question have limited availability and require higher weighting of few individuals. Our results also suggest several more directions for future research. Analyses only weighting for demographics or cardiovascular risk alone had smaller changes than those accounting for both demographics and clinical risk factors, which suggests there may be some interaction between demographic and cardiovascular risk factors worth exploring when analyzing heterogeneous treatment effects, rather than relying only on univariate subgroup analyses (28).

CONCLUSIONS

Overall, our study helps to make quantitative judgments regarding how to generalize trial results to a broader population. Additionally, the methods highlight the specific areas where existing evidence provides the most or least support. Although lack of data in some groups means results supporting intensive BP treatment should be interpreted cautiously, it also highlights areas that need further study to better inform BP treatment guidelines.

ADDRESS FOR CORRESPONDENCE: Dr. Seth A. Berkowitz, Division of General Medicine and Clinical Epidemiology, University of North Carolina School of Medicine, 5034 Old Clinic Building, CB 7110, Chapel Hill, North Carolina 27599. E-mail: seth_berkowitz@med.unc.edu. Twitter: @UNC_SOM, @Stanford.

PERSPECTIVES

COMPETENCY IN MEDICAL KNOWLEDGE: High blood pressure is a major modifiable risk factor for cardiovascular events in patients with diabetes.

TRANSLATIONAL OUTLOOK: Randomized trials should include more diverse cohorts and patients at lower cardiovascular risk to better reflect the target population and enhance clinical practice guidelines.

REFERENCES

1. Fox CS, Golden SH, Anderson C, et al. Update on prevention of cardiovascular disease in adults with type 2 diabetes mellitus in light of recent evidence: a scientific statement from the American Heart Association and the American Diabetes Association. Circulation 2015;132:691–718.

2. Whelton PK, Carey RM, Aronow WS, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/

NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol 2018;71:e127–248.

3. American Diabetes Association. 9. Cardiovascular Disease and Risk Management: Standards of

Medical Care in Diabetes—2018. Diabetes Care 2018;41 Suppl 1:S86–104.

4. Wilt TJ, Kansagara D, Qaseem A, for the Clinical Guidelines Committee of the American College of Physicians. Hypertension limbo: balancing benefits, harms, and patient preferences before we lower the bar on blood pressure. Ann Intern Med 2018;168:369.

5. The SPRINT Research Group. A randomized trial of intensive versus standard blood-pressure control. N Engl J Med 2015;373:2103-16.

6. ACCORD Study Group, Cushman WC, Evans GW, Byington RP, et al. Effects of intensive blood-pressure control in type 2 diabetes mellitus. N Engl J Med 2010;362:1575-85.

7. Bareinboim E, Pearl J. Causal inference and the data-fusion problem. Proc Natl Acad Sci U S A 2016;113:7345-52.

8. Westreich D, Edwards JK, Lesko CR, Stuart E, Cole SR. Transportability of trial results using inverse odds of sampling weights. Am J Epidemiol 2017;186:1010-4.

9. Rudolph KE, van der Laan MJ. Robust estimation of encouragement design intervention effects transported across sites. J R Stat Soc Ser B Stat Methodol 2017;79:1509-25.

10. Rudolph KE, Schmidt NM, Glymour MM, et al. Composition or context: using transportability to understand drivers of site differences in a large-scale housing experiment. Epidemiology 2018; 29:199-206.

11. Najafzadeh M, Schneeweiss S. From trial to target populations: calibrating real-world data. N Engl J Med 2017;376:1203-5.

12. Centers for Disease Control and Prevention. Questionnaires, Datasets, and Related Documentation. National Health and Nutrition Examination Survey. . Available at: https://wwwn.cdc.gov/nchs/nhanes/Default.aspx. Accessed April 24, 2018.

13. Berkowitz SA, Baggett TP, Wexler DJ, Huskey KW, Wee CC. Food insecurity and metabolic control among U.S. adults with diabetes. Diabetes Care 2013;36:3093-9.

14. Berkowitz SA, Meigs JB, Wexler DJ. Age at type 2 diabetes onset and glycaemic control: results from the National Health and Nutrition

Examination Survey (NHANES) 2005-2010. Diabetologia 2013;56:2593-600.

15. Berkowitz SA, Berkowitz TSZ, Meigs JB, Wexler DJ. Trends in food insecurity for adults with cardiometabolic disease in the United States: 2005-2012. PloS One 2017;12:e0179172.

16. Goff DC, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2014;63 Pt B:2935-59.

17. Zhao W, Katzmarzyk PT, Horswell R, Wang Y, Johnson J, Hu G. HbA1c and coronary heart disease risk among diabetic patients. Diabetes Care 2014;37:428-35.

18. Currie CJ, Peters JR, Tynan A, et al. Survival as a function of HbA1c in people with type 2 diabetes: a retrospective cohort study. Lancet 2010;375:481-9.

19. Tillmann T, Vaucher J, Okbay A, et al. Education and coronary heart disease: Mendelian randomisation study. BMJ 2017;358:j3542.

20. Gansevoort RT, Correa-Rotter R, Hemmelgarn BR, et al. Chronic kidney disease and cardiovascular risk: epidemiology, mechanisms, and prevention. Lancet 2013;382:339-52.

21. Polley E, van der Laan MJ. Super learner in prediction. May 2010. UC Berkeley Division of Biostatistics Working Paper Series Working Paper 266. Available at: https://biostats.bepress.com/ucbbiostat/paper266. Accessed July 30, 2018.

22. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. Am J Epidemiol 2008;168:656-64.

23. Lee BK, Lessler J, Stuart EA. Weight trimming and propensity score weighting. PloS One 2011;6: e18174.

24. Blood Pressure-Lowering treatment Trialists Collaborators. Blood pressure-lowering treatment based on cardiovascular risk: a meta-analysis of individual patient data. Lancet 2014; 384:591-8.

25. Basu S, Sussman JB, Rigdon J, Steimle L, Denton BT, Hayward RA. Benefit and harm of intensive blood pressure treatment: Derivation and validation of risk models using data from the SPRINT and ACCORD trials. PLoS Med 2017;14: e1002410.

26. Buckley LF, Dixon DL, Wohlford GF, Wijesinghe DS, Baker WL, Tassell BWV. Intensive versus standard blood pressure control in SPRINT-eligible participants of the ACCORD-BP trial. Diabetes Care 2017:dc171366.

27. Cholesterol Treatment Trialists' (CTT) Collaborators, Mihaylova B, Emberson J, Blackwell L, et al. The effects of lowering LDL cholesterol with statin therapy in people at low risk of vascular disease: meta-analysis of individual data from 27 randomised trials. Lancet 2012;380:581-90.

28. Hayward RA, Kent DM, Vijan S, Hofer TP. Multivariable risk prediction can greatly enhance the statistical power of clinical trial subgroup analysis. BMC Med Res Methodol 2006;6:18.

KEY WORDS diabetes mellitus, generalizability, hypertension, transportability

APPENDIX For an expanded Methods section as well as supplemental figures and tables, please see the online version of this paper.