TITLE: Blood pressure and cognition: Factors that may account for their inconsistent association

AUTHORS AND AFFILIATIONS: Melinda C. Power[1,2]; Eric J. Tchetgen Tchetgen[3]; David Sparrow[4,5]; Joel Schwartz[1,2]; Marc G Weisskopf[1,2]

[1]Department of Epidemiology, Harvard School of Public Health, Boston MA USA
[2]Department of Environmental Health, Harvard School of Public Health, Boston MA USA
[3]Department of Biostatistics, Harvard School of Public Health, Boston MA USA
[4]Veterans Affairs Normative Aging Study, Boston MA USA
[5]Department of Medicine, Boston University School of Medicine, Boston MA USA

eAppendix

Details of our marginal structural models.

We used the following marginal structural models for our primary analyses. For our primary analyses, both were weighted using a single set of inverse probability weights that combine three different inverse probability weights – one for confounding, one for censoring due to death, and one for censoring due to non-death drop-out.

(A) $E(Y_{\bar{a}}(k)) = \beta_0 + \beta_1 I(a(k) = 1) + \beta_2 k$

(B) $E(Y_{\bar{a}}(k)) = \beta_0 + \beta_1 I(a(k) = 1) + \beta_1 I(a(k) = 1) * (cum(\bar{a})) + \beta_1 I(a(k) = 1) * (k - cum(\bar{a})) + \beta_4 k$

Where k is the age at cognitive testing

$I(a(k)=1)$ is an indicator for whether the participant has a history of hypertension at age k

$cum(\bar{a})$ is a cumulative measure of years since hypertension initiation

$k\text{-}cum(\bar{a})$ is the age at initiation of hypertension

We refer the reader to a paper by Hernan and colleagues[1] for sample code and a discussion of the derivation and use of stabilized inverse probability weights in the context of longitudinal data. However, we reproduce formulas for inverse probability of exposure weights and inverse probability of censoring weights here, adapted to the context of our study, and discuss their application in the current analyses.

Details of computation of stabilized inverse probability weights.

We calculated stabilized inverse probability of exposure weights, which address confounding, in our longitudinal data using the following process:

(a) We began with a dataset with one row per study visit. We used a logistic regression model to model the relationship between converting to having a history of hypertension at the next visit given the history of hypertension status at the current visit, the history of time-varying confounders, and the baseline covariates. Note that history of hypertension is monotonic, such that once a participant converts to having a history of hypertension the probability of having a history of hypertension at the next visit is one. As such, our dataset for the logistic regression model only contained visits at which a person has not yet converted to having a history of hypertension. By extension, history of hypertension was not included as a variable in the logistic model because there is no variation in the data – every visit in the dataset has no history of hypertension at the current visit, although they could either remain non-hypertensive or become

hypertensive at the next visit. We then created a variable with the predicted probability of converting to having a history of hypertension at the next visit from the logistic regression model for all visits where the participant had not yet converted to having a history of hypertension and added it to the full dataset with one line per visit. This variable was used to derive the denominator of our stabilized inverse probability of exposure weights.

(b) We then repeated this process, omitting all covariates from the logistic regression model. This second variable was used to derive the numerator of the stabilized inverse probability of exposure weights.

(c) We used the following formula to calculate the stabilized inverse probability of exposure weight for each visit (indexed by t) summarized in eTable 1.

$$\prod_{t=1}^{t} \frac{\Pr(Ht_t \mid \overline{H}t_{t-1})}{\Pr(Ht_t \mid \overline{H}t_{t-1}, \overline{L}_t, V)}$$

Ht_t	= History of hypertension at visit t
$\overline{H}t_{t-1}$	= History of hypertension prior to visit t
$\overline{L}_t$	= History of time-varying covariates up through visit t
V	= Baseline covariates

Note that the probabilities in the formula do not refer to the predicted probability of having a history of hypertension at visit t, $Pr(Ht_t=1)$, or the predicted probability of lacking a history of hypertension at visit t, $Pr(Ht_t=0)$. Instead, the probability of interest is the predicted probability

of having the history of hypertension status actually observed at visit t. Also note that the quantity following the product term was one for all visits where participants had already converted to having a history of hypertension at visit t-1.

(d) In our case, we use a dataset with one row per person for our marginal structural linear model because we have one outcome measure per person and use summary exposure variables. Therefore, the inverse probability of exposure weights from the final visit of each participant (t=max) as calculated above are the final weights we would use in our analyses if we were not also using weights for dependent censoring.

Stabilized inverse probability of censoring weights, which account for dependent censoring, can be calculated using a similar process. Because we hypothesized two different censoring mechanisms, we actually calculated two sets of weights. For weights to account for censoring due to death, we used the following formula:

$$\prod_{t=1}^{t} \frac{\Pr(D_t = 0 \mid \overline{Ht}_{t-1})}{\Pr(D_t = 0 \mid \overline{Ht}_{t-1}, \overline{L}_t, V)}$$

D_t = Died prior to visit t? (1=yes, 0=no)

$\overline{Ht}_{t-1}$ = History of hypertension prior to visit t

$\overline{L}_t$ = History of time-varying covariates up to visit t

V = Baseline covariates

And for censoring due to non-death drop-out, we used the following formula:

$$\prod_{t=1}^{t} \frac{\Pr(C_t = 0 \mid D_t = 0, \overline{Ht}_{t-1})}{\Pr(C_t = 0 \mid D_t = 0, \overline{Ht}_{t-1}, \overline{L}_t, V)}$$

C_t = Non-death drop-out prior to visit t? (1=yes, 0=no)

D_t = Died prior to visit t? (1=yes, 0=no)

$\overline{Ht}_{t-1}$ = History of hypertension prior to visit t

$\overline{L}_t$ = History of time-varying covariates up to visit t

V = Baseline covariates

As before, we began with a dataset with one line per study visit. It is important to note that the variables that predict drop-out (L,V) may differ by censoring mechanism (death, non-death). It is also important to note that our censoring mechanisms are ordered in time, with death preceding non-death drop-out, which accounts for the inclusion of D_t=0 in the denominator of the weights for non-death censoring.

The process for the calculation of each set of inverse probability of censoring weights is similar. For each set of weights, we created a dataset including all visits, with one line per visit, where the person could conceivably be censored by the mechanism of interest prior to the next visit. The dataset for censoring due to death contained all visits except those where the person completes cognitive testing, as we administratively censor each individual at that point. The dataset for censoring due to non-death drop-out contained all visits except those where the

person completes cognitive testing or is known to die prior to the next potential visit. We used logistic regression models to predict the probability of not being censored at the next visit and used these predicted probabilities to calculate the weights using the formulas above. Again, for the current analyses, since our analytical dataset has only one line per person, if we were only applying a single weight for censoring, we would weight our analyses by the stabilized inverse probability of censoring weight at the last study visit. This inverse probability of censoring weight is the product of the final inverse probability of censoring weights for death and the final inverse probability of censoring weights for non-death drop-out.

In our study, we needed a single set of weights that function as inverse probability weights for confounding, censoring due to death, and non-death censoring. To compute these summary weights, we take the inverse probability of exposure weight, the inverse probability of censoring due to death weight and the inverse probability of censoring due to drop out weight at the participant's last visit and simply multiply them together. Characteristics of each set of inverse probability weights calculated from our data and the final combined weight used in our analyses are provided in eTable 1.

References

1. Hernan MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. Epidemiology. 2000 Sep;11(5):561-70.

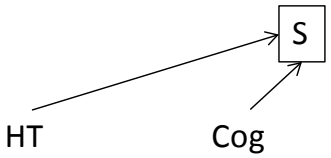
eTable 1. Characteristics of Final Stabilized Inverse Probability Weights Among Persons Who Completed Cognitive Testing

	Weight			
	Censoring Due To Death	Censoring Due to Drop-Out	Confounding	Combined Weight
Mean (Standard Deviation)	0.99 (0.36)	1.02 (0.38)	1.01 (1.82)	1.00 (1.85)
Minimum	0.62	0.82	0.10	0.10
1st Percentile	0.68	0.84	0.12	0.12
5th Percentile	0.74	0.86	0.23	0.22
10th Percentile	0.77	0.88	0.30	0.27
50th Percentile	0.91	0.96	0.67	0.62
90th Percentile	1.24	1.15	1.68	1.67
95th Percentile	1.47	1.33	2.42	2.61
99th Percentile	2.52	1.59	6.70	8.72
Maximum	5.07	10.58	38.47	27.36

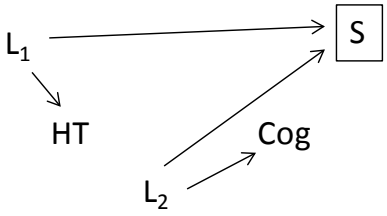
eTable 2. Sensitivity analyses for the association between a 1 year difference in age at onset and duration since hypertension initiation prior to cognitive testing and age-adjusted mean cognitive test z-score.

	Beta (95% Confidence Interval)			
	Alternate derivation of age at onset of hypertension	Restricting to those under 35 at enrollment	Incorporating a 4 year prodromal period	Truncating weights at the 1st/99th percentile
Age at onset	-0.01 (-0.03, 0.004)	0.001 (-0.03, 0.04)	-0.006 (-0.02, 0.01)	-0.002 (-0.02, 0.02)
Duration Since Initiation	-0.03 (-0.05, -0.01)	-0.02 (-0.05, 0.02)	-0.02 (-0.04, 0.002)	-0.02 (-0.03, 0.003)

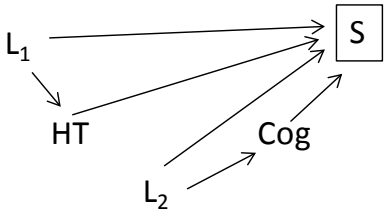
A)



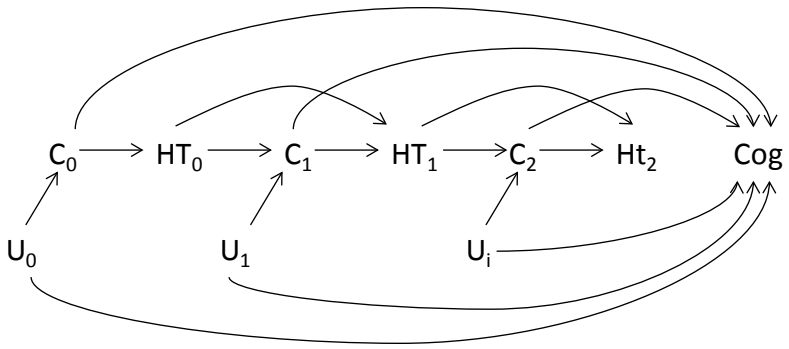
B)



C)



eFigure 1. Simple causal direct acyclic graphs under the null hypothesis of no association between hypertension and cognition that show causal structures resulting in selection bias due to association between study participation (S) and both hypertension (HT) and cognitive function (Cog). In A) hypertension and cognitive function both have direct effects on participation. In B) the association between participation and both cognition or hypertension attributable to a common cause (L). For example, L could be a genetic variant or lifestyle factors, like diet. If the entire association between participation and both hypertension and cognition were due to L, and if we were able to measure and adjust for L in our regression models, there would be no bias. However, the causal structure is more likely a combination of A) and B), as illustrated in C).



eFigure 2. A simple, explanatory causal direct acyclic graph under the null hypothesis of no association between hypertension and cognition and the assumption of no unmeasured confounding illustrating the potential for time-varying confounding in our data (although we show only three study visits prior to cognitive assessment here, the pattern can easily extend to any number of assessments). HT_i is hypertension diagnosis at time i, Cog is cognitive test scores at the end of follow-up, C_i is a time-varying confounder at time i, and U_i is an unmeasured, time-varying common cause of C and Cog. Note that C meets both criteria for time-varying confounding, as it A) predicts both cognitive status at the end of follow-up and subsequent hypertension diagnosis and B) is predicted by prior hypertension diagnosis. Associations obtained without conditioning on C are biased, as C is a common cause of hypertension diagnosis and cognition, and associations obtained with conditioning on C are biased due to the backdoor path created due to conditioning on a collider (C), suggesting the need for other methods (e.g. inverse probability weighting) to obtain unbiased estimates.