



Of means standard deviations and blood pressure trials

CHRISTOS ARGYROPOULOS¹

1. University of New Mexico

READ REVIEWS

WRITE A REVIEW

CORRESPONDENCE:
argchris@hotmail.com

DATE RECEIVED:
June 10, 2015

KEYWORDS:
hypertension guidelines,
descriptive statistics

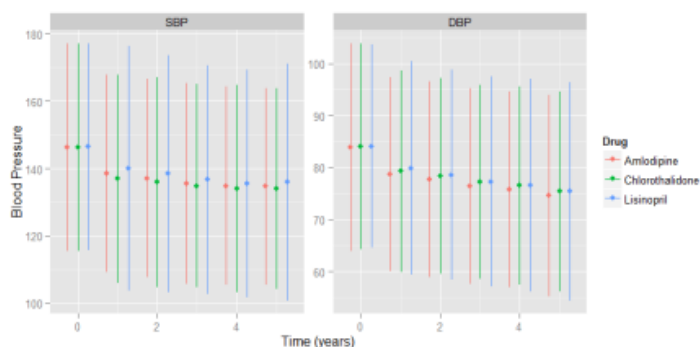
© Argyropoulos This article is distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and redistribution in any medium, provided that the original author and source are credited.



Last week, the long awaited Joint National Commission (JNC) released their 8th revision of their [hypertension management guidelines](#) (more than 10 years after the 7th version!). JNC 8 will join the plethora of guidelines ([ASH/ISH](#), [ESH](#), [KDIGO](#) and the [ACC/AC advisory](#) that will be upgraded to full guideline set later on) that may be used (or not) to justify (or defend) treating (or not treating) patients with elevated blood pressure (BP) down to pre-approved, committee blessed targets. Though it is not my intention to go into the guidelines (since opinions are like GI tracts: everyone has one, whereas biases are like kidneys: everyone has at least one and usually two), it is interesting to seize this opportunity and revisit one of the largest blood pressure trials that is used to defend certain aspects of the guidelines recommendations.

The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial ([ALLHAT](#)) is one of the largest hypertension trials and was all the rage 12 years ago when failing to meet its primary endpoint i.e. the superiority of various first line antihypertensive agents, it reported differences in various secondary outcomes (including blood pressure control) favoring the oldest (chlorothalidone) v.s. the newest (lisinopril/amlodipine) medication. Since 2002 all agents used in ALLHAT have gone generic (and thus have the same direct costs), diluting the economic rationale for choosing one therapy v.s. the other. Nevertheless the literature reverberated for years ([2003](#) , [2007](#) v.s. [2009](#)) with intense debates about what ALLHAT showed (or didn't show). In particular the question of the blood pressure differences observed among the treatment arms was invoked as one of the explanations for the diverging secondary outcomes in ALLHAT. In its simplest form this argument says that the differences of 0.8 (amlodipine)/2 (lisinopril) mmHg of systolic BP over chlorothalidone may account for some of the differences in outcomes.

I was exposed to this argument when I was an internal medicine resident (more than a decade ago) and it did not really make a lot of sense to me: could I be raising the risk for cardiovascular disease, stroke, and heart failure by 10-19% of the patient I was about to see in my next clinic appointment by not going the extra mile to bring his blood pressure by 2 mmHg? Was my medical license, like the 00 in 007's designation, a license to kill (rather than heal)? Even after others explained that the figures reported in the paper referred to the **difference in the average blood pressure between the treatment arms**, not the difference of blood pressure within each patient I could not get to my self to understand the argument. So more than a decade after the publication of ALLHAT I revisited the paper and tried to infer what the actual blood pressures may have had been (no access to individual blood pressure data) during the five years of the study. The relevant information may be found in Figure 3 (with a hidden bonus in Figure 1) of the [JAMA paper](#) which are reproduced in a slightly different form below (for copyright reasons):



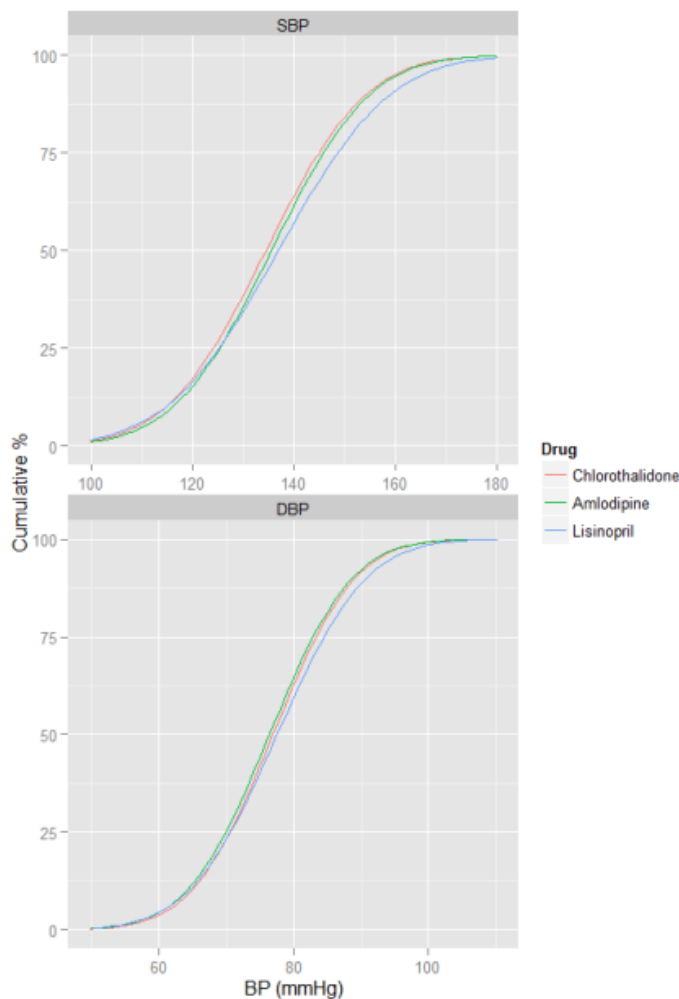
The figure shows the Systolic (SBP) and Diastolic (DBP) pressures of the participants who continued to be followed up during the ALLHAT study; unsurprisingly for a randomized study there were almost no differences before the study (T=0). Differences did emerge during follow-up and in fact the mean/average ("dot") in the graph was lower for chlorothalidone v.s. both amlodipine and lisinopril during the study. However, the standard deviation (the dispersion of the individuals receiving the study drugs around the mean: shown as the height of vertical line) was also larger (by about 17%) for lisinopril suggesting that there both more people with **higher** and **lower** blood pressures compared to chlorothalidone.

This pattern is also evident if one bothers to take a look at the flowchart of the study (Figure 1) which lists patterns/reasons for discontinuation during the 5th year. Among the many reasons, "Blood Pressure Elevation" and "Blood Pressure Too Low" are listed:

A garden variety logistic regression for the odds of discontinuation shows there is no difference between amlodipine and chlorothalidone due to high ($p=0.16$) or low ($p=0.74$) BP. On the other hand, the comparison between lisinopril and chlorothalidone is more interesting:

- Odds of **discontinuation due to low BP**: 1.44 95%CI: 1.00-2.06 ($p=0.044$)
- Odds of **discontinuation due to high BP**: 3.38 95%CI: 2.35-4.87 ($p<0.001$)

So despite the higher mean, the higher standard deviation implies that there more patients with low (and too low) BP among the recipients of lisinopril. In other words, blood pressure control was more variable with lisinopril compared to the other two drugs: this variability can be quantified by using the reported means/standard deviations to look at the cumulative percentage of patients with BP below a given cutoff (for simplicity we base the calculations on the year 3 data):



So it appears that lisinopril (at least as used in ALLHAT) was able to control more patients to < 120/70 (which are low levels based on the current guidelines), but fewer patients at the higher end of the BP spectrum. The clinical translation of this observation is that there will be patients who will be ideal candidates for lisinopril (maybe to the point where dose reduction is necessary) and others who will fail to respond, so that individualization of therapy, rather than one size fits all is warranted. Such individualization may be achieved either on the basis of [short shared physician/patient decision making, n of 1 trials](#), biomarker levels (e.g. home blood pressure measurements) or demographic profiling (as is done in JNC8 for African American patients).

Notwithstanding these comments, one is left scratching one's head with the following questions:

- who were the patients with an exaggerated and dampened out response to lisinopril in ALLHAT
- could the variability in BP control provide a much better explanation for the variability in secondary outcomes in ALLHAT? (the investigators did apply what is known as a time-updated analysis using the observed BPs during the trial, but this is not the statistically proper way to analyze this effect in the presence of loss-to-follow up and possibly informative censoring)
- what are the clinical implications of lowering BP to a given level when this is done with different classes of agents? This question is related to the previous one and both are not unanswerable with time-updated models of endogenous (such as BP readings) variables

At a more abstract level, should we scrutinize paper tables for the means as well as the standard deviations of response variables looking for hidden patterns that may not be evident at a first look? In the clinic one is impressed with the variability of the patient responses to interventions, yet this variability is passed over when analyzing, reporting and discussing trial results in which we only look at the means. To me this seems a rather deep inconsistency between our behaviours and discourse with our Clinician v.s. our Evidence Based Medicine hats on, which may even decrease the chances of finding

efficacious and effective treatments. Last but certainly not least, how can we begin to acknowledge variability in trial design, execution, analysis and reporting so as to better reflect what actually happens in the physical world, rather than the ivory towers of our statistical simulators?