Author's response to reviews

Title:A Clinical Prediction Score for Diagnosing Unilateral Primary Aldosteronism May Not Be Generalizable

Authors:

Erik S Venos (erik.venos@gmail.com)
Benny L So (benny.so@albertahealthservices.ca)
Valerian C Dias (valerian.dias@cls.ab.ca)
Adrian Harvey (adrian.harvey@albertahealthservices.ca)
Janice Pasieka (janice.pasieka@albertahealthservices.ca)
Gregory A Kline (gregory.kline@albertahealthservices.ca)

Version: 2
Date: 22 November 2014

Author's response to reviews: see over
Response to reviewers

Reviewer: Gian Paolo GP Rossi
Reviewer's report:
Kline et al performed a single-centre retrospective study to confirm the result of a recent study by Kupers et al that proposed a clinical prediction score to select the patients to be submitted to AVS. The conclusions of this study are that this score has a low specificity. While of interest this conclusion could be anticipated given that the score was based on CT (which, by definition, provides only imaging data and no functional information), and on the presence of hypokalemia and/or hyperfiltration.
Nonetheless, the results are important and could be of interest for the Journal. However, the following issues deserve attention.

Reviewer #1: Thank you for these suggestions; we completely agree that the original score was destined to be less than ideal based upon its dependence on CT findings. Hopefully our paper shows the value of external validation exercises and yet respectfully recognizes the role that a prediction rule might have in helping (not replacing) the decision to use adrenal vein sampling.

Major criticisms
1. The last sentence of the abstract is rather obscure: what does it mean “local validation”? Either a score is valid and generalizable or it is not.. Please rephrase

Agreed, we have re-phrased as:

“At best, a high score in this prediction rule may be an additional tool for helping to confirm a decision to offer patients adrenal vein sampling.”

2. Incorporation of hyperfiltration in the score is a naïve assumption, because PA patients have hyperfiltration initially, but develop hypofiltration when they develop renal damage. Hence, using hyperfiltration in a score without taking into account the known duration of hypertension and the CKD class is probably too simplistic. This should be discussed.

Thank you, we agree this is an excellent point. We have now added this to the discussion:

The clinical prediction score itself has several important limitations to widespread use, notably the use of GFR as a predictive value. A person with PA may have widely variable measures in GFR depending upon age, concomitant medication use and duration of aldosteronism, all of which can influence a single measure of GFR without reference to the laterality status.

3. The gold standard to validate any diagnostic score or test should be outcome. In this regards it is not entirely clear what was done for the ROC analysis that requires a binary outcome. Moreover, using only the fall in BP after
adrenalectomy is not per se sufficient to confirm the diagnosis of lateralized PA as shown by Gordon’s group. The Authors should provide the ARR data post-op and also give data on normalization of PRA, aldosterone and the ARR, or at least acknowledge this as a limitation of their study if data are unavailable.

Thank you, we completely agree that fall in BP post operatively does not confirm a diagnosis of PA but rather one confirms the diagnosis through biochemistry, pathology and clinical features (and possibly immunohistochemistry although not available here).

In the patient results section, we have included data of all of the above variables and we acknowledge that less than 100% of our patients follow through with repeat ARR measurements:

“On surgical pathology, 34 had discrete adrenocortical adenoma, 14 had discrete adenoma with surrounding adrenal hyperplasia, and 5 had pure unilateral adrenal hyperplasia. Of the 37 patients who underwent postoperative ARR, 36 had an undetectable PAC or a normal ratio (ARR <550). The one patient with a persistently high post-operative ARR had an adrenocortical adenoma on pathology while also having a nodule on the contralateral adrenal gland, presumably representing a second aldosterone producing adenoma and an initial false positive AVS. Excluding the 3 patients lost to long term follow-up, 23 patients (44%) achieved clinical hypertension cure and sustained BP control of less than 140/90 mmHg and 28 patients (54%) achieved a marked improvement with a reduction of medications and control of BP < 140/90 mmHg.”

We now include the following admission in our discussion in order to be clear about the limitations in study data:

“Ideally all subjects should have had a post-operative ARR measured to confirm resolution of PA rather than depending upon blood pressure and pathology alone.”

5. Discussion is too long and wordy. Some sections can be deleted.

OK, thank you. We’ve gone through and tried to clean this up removing unnecessary numbers and phrases, hopefully to make it easier and more succinct to read.
Minor comments
Page 4 line 11 what is a high success rate? Please give a number.
At the time of database censor for this study, our AVS success rate was 97%, now included in the methods section.

Line 17 what is seated, upright? Please specify.
Blood is drawn with the patient in a seated, upright position. We have now clarified this in the methods.

Line 19 It is usually advised to use 0.2 ng/ml/h to avoid inflation of the ARR. Why 0.1 was used? This caused a two-fold inflation.

Thank you for this comment – the reviewer is right that there seems to be a lot of variability among different investigators in terms of whether one uses an analytical sensitivity vs a limit of detection term for “correcting” an inflated ARR. For PRA, we used a similar Diasorin PRA RIA as did Kupers et al and Candy Sze et al.

According to the package insert, the analytical sensitivity is 0.018 ng, so, 0.018 ng/0.1 mL (sample volume) = 0.18 ng/mL

0.18 ng/mL * 1.11 mL/ 1.00 mL (dilution factor) = 0.20 ng/mL

0.20 ng/mL / 1.5 h (angiotensin I generation time) = 0.13 ng/mL/h

For simplicity’s sake we have therefore adopted a lower reporting limit of 0.1 ng/ml/h based on these results.

Page 11 line 24 CPS are an important tool. Please correct to … is …. Or remove .. an …. Thank you, we have re-written it for better clarity.

Figure 2: please report 95% CI and Youden index on the ROC curve. The font size of axis labels is too small.

Agreed, thank you. We have re-drawn the figure and added the requested data.

Level of interest: An article of importance in its field
Quality of written English: Acceptable
Statistical review: Yes, and I have assessed the statistics in my report.
Declaration of competing interests: I declare that I have no competing interests.
The manuscript of Venon et al proposes the validation of a clinical prediction score for the diagnosis of unilateral primary aldosteronism. They have demonstrated that this prediction score, which presented 100% of specificity in the original cohort, presented lower specificity in the Canadian cohort of PA. The study is technically well done, well written and the size of the cohort is adequate. Some points, however, must be clarified or discussed.

Major compulsory revision

1- Previous studies have already demonstrated that the prediction score of Kupers et al did not present 100% of specificity in different cohorts of primary aldosteronism. What are the contributions of the present study for the clinical management of patients with PA? This information should be added in the manuscript.

Thank you, this is a very fair question. Most of the prior studies (and ours, as well) have reported on the overall use of the prediction score, which turns out to be quite limited. However, we draw attention to the small number of patients in whom a very high score does appear to have excellent specificity, adding to the clinical confidence in the diagnosis. This point is now further clarified in the discussion:

A CPS of \( \geq 6 \) did perform better with a specificity (96%) approaching 100%. If this cut point were used, then the CPS would apply to approximately 20% of patients with PA. This point has not been previously emphasized; although few patients may achieve such a high score, for those who do, a near-perfect specificity may well add significantly to the clinical confidence in a lateralized diagnosis.

2- Recent studies have shown different population characteristics among PA cohorts despite recent guidelines for the management of primary aldosteronism. This could be reinforced in the discussion, with references of different cohorts as well as hypothesis for these cohort differences. Do the authors think that differences in ethnical background or differences in procedures of recruitment could influence in the differences observed among the different studies?

Please see answer to question #3 which is similar.

3- The scope of the present study is based in clinical and biochemical approaches for PA diagnosis. However, recent APA genetic advances must be
included in the discussion section, maybe as a hypothesis for the difference of clinical score specificity across the different studies. In the present study PA is largely more frequent in females than males (table 2) and is known that somatic mutations in KCNJ5 are also more frequent in females. Could this represent a genetically bias between the derivation cohort and the cohort described in this manuscript?

These are very interesting speculations raised by the reviewer. I think the whole PA field is now recognizing that tumour/hyperplasia biology is non-uniform across subjects and thus variability in etiology, presentation and therapy response is to be expected. Genetic, ethnic and sex differences may be very important predictors of these different biologies and the exact contribution of each is still being determined.

We are happy to add a section and references speculating as to these potential population differences:

It is increasingly recognized that PA can have multiple different molecular etiologies including various genes and ectopic hormone receptors (25,26). It would therefore be expected that there may well be significant variation in disease presentation, diagnostic features and possibly response to therapy according to such differing causes and modified by variables that influence penetrance such as age, sex and ethnicity. In the future, more sophisticated molecular studies of confirmed PA cases may permit a better understanding of the apparent differences between PA study populations.

4- In the discussion section, page 10, the authors suggested that the differences observed on AVS data among the present manuscript and the derivation cohort “may reflect differing definitions of a unilateral AVS or different definition of an adrenal nodule”. This sentence is not clear. What are the “definitions of unilateral AVS or adrenal nodules” in the different cohorts? This point needs to be clarified.

Thank you, we have re-written as:

This may reflect differing biochemical definitions as to what constitutes “unilateral” AVS or a differing size definition of an adrenal nodule. Centers using more strict or more lenient lateralization indices may well show apparent differences in the proportions of patients deemed unilateral and verification will be limited to those who actually undergo adrenalectomy.

5- Are there other clinical scores for the prediction of unilateral PA recently
described? A description of these scores for comparison with the score used should be added in the discussion section.

Thank you – yes, there is now at least 1 other CPS published. We now mention these in the discussion with a compare and contrast approach along with an overview of future needed directions in prediction score development.

A second PA subtype prediction rule has also recently appeared in the literature (29) which purports to show that the combination of plasma aldosterone, potassium and post-captopril ARR may help predict unilaterality. However, this has also been generated using a small sample size and will also need external validation attempts. A multi-national standardized PA database is likely necessary to have adequate power for the discovery of all significant clinical factors that may ultimately form a more accurate and generalizable clinical prediction rule.

Minor Essential revision
1- Table 1. This table reflects the score used in this manuscript but it was already proposed and described in a previous study. A reference of the original study must be added in this table.

Agreed, we have now added the reference.

Level of interest: An article of importance in its field
Quality of written English: Acceptable
Statistical review: No, the manuscript does not need to be seen by a statistician.
Declaration of competing interests: I declare that I have no competing interests