

Supporting Elaboration

Leave-one-out Analysis

In the text we speak of what we believe to be a genuine four-factor interaction in split three of the tree grown on data from Chinese women. Readers will see that this is an example for which Model A seems to apply. We give first a table with Hotelling's T^2 and permutation/bootstrap χ^2 p values for the significance of the four variables involved in the split. Note that none is close to 0.10, let alone 0.05. However, Permutation t for the reduction in impurity for this split is 4.739, a highly significant value judged by the nominal standard normal sampling distribution that applies, at least approximately. While these p values are not conditional on the computation at split one, we believe the signal in the various analyses is strong enough that conditioning on the previous split—something we do not know how to do in any case—would be irrelevant.

Table 6. Chinese women, split three (SNP for non-wild-type)

| Variable | T^2 p value | χ^2 test p value |
|---------------------------|-----------------|-------------------------|
| Insulin resistance (+, R) | 0.679 | 0.141 |
| AGT2R1A1166C (R) | 0.702 | 0.211 |
| AVPR2G12E (P) | 0.759 | 0.797 |
| PTPN1x9INV (R) | 0.930 | 0.460 |
| Permutation $t = 4.739$ | | |

The next table gives the learning sample reduction in the impurity of the cited split two, first for the four chosen predictors, and then in order with each of the four backed out of the computation. The next column gives comparable results, with means and standard deviations, for fivefold cross-validation done 50 times in each case. Note that in each case the Permutation t is less for the three variables model than for the model with all four variables, the difference being especially striking when insulin resistance status is deleted.

Table 7. Leave-one-out performance

| | Learning sample risk | fivefold cv 50 time avg (SD) | Perm t |
|-----------------|----------------------|------------------------------|----------|
| All four | 0.026 | 0.016 (0.005) | 4.739 |
| Out insulin | 0.012 | 0.002 (0.003) | 1.229 |
| Out <i>AVPR</i> | 0.022 | 0.013 (0.004) | 3.597 |
| Out <i>PTP</i> | 0.017 | 0.006 (0.003) | 3.819 |
| Out <i>AGT</i> | 0.017 | 0.007 (0.004) | 3.173 |

Because the mean values are correlated, the p values were done again, accounting for the computed covariance structures between the four-predictor model and the successive three-predictor models. In this scenario we computed p values for the difference between the four-variable model and the successive three-variable models. When insulin resistance status is deleted, the p value for the difference is 0.012; when *AVPR* (arginine vasopressin receptor two) is deleted, the p value for the difference is 0.420; with *PTP* (protein tyrosine phosphatase 1B) left out, $p=0.034$; with *AGT* (angiotensin II type one receptor) out, the p value is 0.060. Note the consistency of patterns, these p values and Permutation t values in Table 7.

Results on Japanese

Before we studied the data carefully we thought that the impact of our SNP genotypes, if they bear upon hypertensive status, should predispose Chinese and Japanese women in the same way. We have described difficulties with the SAPHIRE sampling scheme earlier, but return here to emphasize these differences. Recruitment in Taiwan focused far more upon hypotensive sibs of hypertensive probands than did recruitment in Hawaii. Of course, the majority of our Chinese women were from Taiwan and the majority of Japanese women were from Hawaii. No matter the “true” fractions of hypertensives and hypotensives within families in these two groups, the prevalence of hypertensives in our sample was far higher for the Japanese than for the Chinese. There are in total 161 samples, 23 hypotensive and 138 hypertensive, which is unlikely to be the distribution in general population. Moreover, classifying with the same products of priors and misclassification costs (2:1 in favor of hypotensive in the Chinese) produced

nothing of interest in the smaller Japanese group. When the ratio was changed to 6:1, a story emerged: the final tree has two splits.

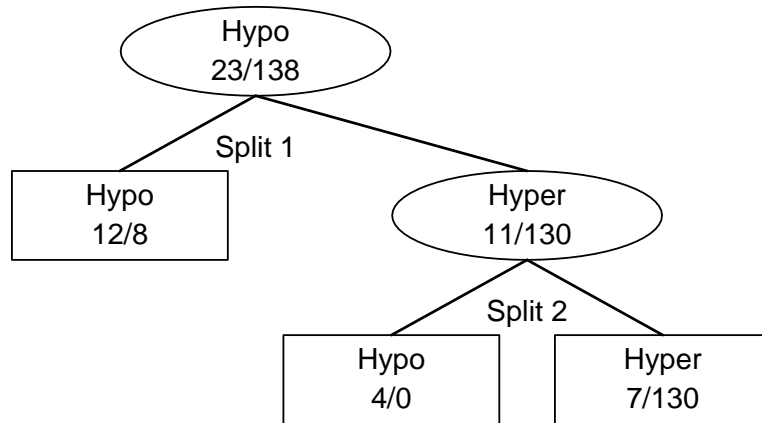


Fig. 2. The ovals and rectangles, respectively, indicate internal and terminal nodes. The label assigned to each node is determined so as to minimize the misclassification cost. The number to the left of the “/” is the number of hypotensives; the number to the right is the number of hypertensives. Here we assume equal prior probabilities and misclassification cost for hypotensives twice that for hypertensives.

The detailed information on the two splits are summarized in the following two tables:

Table 8. Japanese women, split one

| Variable | T^2 p value | χ^2 p value |
|-------------------------|-----------------|--------------------|
| AGT2R2C1333T (R) | 2.66e-05 | 0.263 |
| MLRI2V (P) | 6.96e-05 | 0.301 |
| Cyp11B2.5.aINV (R) | 0.011 | 0.406 |
| Cyp11B2x1INV | 0.131 | 0.175 |
| Permutation $t = 5.860$ | | |

Table 9. Japanese women, split two

| Variable | T^2 p value | χ^2 p value |
|-------------------------------|-----------------|--------------------|
| BADG16R (P) | 0.0025 | 0.714 |
| AGT2R2C1333T (R) | 0.0086 | 0.247 |
| KLKQ3E (2/2 R, 1/2 P) | 0.0249 | 0.812 |
| Cyp11B2.5.aINV (2/2 P, 1/2 R) | 0.0418 | 0.917 |
| PTPN1i4 (R) | 0.681 | 0.401 |
| Permutation $t = 4.457$ | | |

Although what it suggests is unclear at this writing, the performance on learning sample and cross-validation is described in the table below.

Table 10. FlexTree performance on Japanese women

| Learning sample | | | |
|-----------------------|-------|------|---|
| True \ Classification | Hyper | Hypo | $Sensitivity = \frac{130}{138} = 0.942$ |
| Hyper | 130 | 8 | $Specificity = \frac{16}{23} = 0.696$ |
| Hypo | 16 | 7 | $Miscost = 8 + 7 \times 6 = 50$ |
| Cross validation | | | |
| True \ Classification | Hyper | Hypo | $Sensitivity = \frac{122}{138} = 0.884$ |
| Hyper | 122 | 16 | $Specificity = \frac{9}{23} = 0.391$ |
| Hypo | 14 | 9 | $Miscost = 16 + 6 \times 14 = 100$ |

Note that classification was reasonably successful, with cross-validation showing 30 misclassified among 161. Given that the Japanese were older than the

Chinese, it is not surprising that by and large the SNPs that figure in our two-split tree are nearly all SNPs that figure in the split of postmenopausal Chinese women. As one might expect, age did not figure in either split for Japanese. As with the Chinese, there was a wide difference between significance as judged by Hotelling's T^2 and by bootstrap/permutation χ^2 , implying that any impact of genes is additive.