Index

Note: An f following a page number refers to a figure on that page; a t following a page number refers to a table following that page.

adaptive
eligibility, 55–57
randomization, 31
adaptive signature design, 72–73, 74
cross-validated, 75–79
fall-back test, 47–48
adaptive threshold design, 59–63
probabilistic indication classifier and, 60–62
sample size planning, 62–63
aggregated Classification Trees, 117
algorithms
class prediction, 112–118
genetic, 114
pre-specified algorithmic analysis plan, 72
analytical validity, 11, 88
baseline cumulative hazard, 122–123
baseline hazard function, 30, 120
Battle I trial in advanced non-small cell lung cancer, 30–33
results of, 31, 31–32
two-stage design, 32–33
Bayesian methods, 101–103
adaptive designs for randomized phase II trials, 30–33
computing posterior probabilities, 102–103
estimating sensitivity and specificity, 124
frequentist methods versus, 101–102
loss function, 102–103
non-informative prior distributions, 103
posterior distributions, 102
prior distributions, 102, 103
probabilistic indication classifier, 61
Type I error, 103
usefulness in phase 3 trials, 103
binary disease classification, 124–125
biomarkers. See predictive biomarkers;
prognostic biomarkers
blinding
assay to clinical data, 87, 88
results of interim analyses, 5
bootstrapping, 61, 66, 79
BRB-ArrayTools software, 125
breast cancer
enrichment design and trastuzumab, 39
MammaPrint score, 11, 22, 33
Oncotype DX recurrence score, 11, 22–23, 35
TAILORx clinical trial, 22–23, 35, 86
calibration, 119–120
class comparison, 105–106, 107
class discovery, 105–106
class labels, 107–108, 109, 110, 111
class prediction, 105
bias in estimate of error rates, 111
cross-validation, 68, 109–111
leave-one-out-cross-validation, 109–111
definition of, 107–108
estimating accuracy of, 108–112
feature selection, 108, 113–114
mathematical classifier function, 108, 112–118
classification trees, 117
compound covariate, 70, 116
Fisher discriminant analysis, 115, 116
k-nearest neighbor, 117
nearest centroid classification, 117
nearest neighbor classification, 116, 117–118, 125
support vector machines, 116
weighted voting classifier, 69, 116
misclassification rate, 116, 118
parameter specification, 108
clinical trials. See phase 1 trials; phase 2 trials; phase 3 trials
clinical validity, 11, 85
cluster analysis, 106
companion diagnostic, 35–36, 88
Compound Covariate Predictor, 116
confidence interval, 66, 94–95
covariance matrix, 71, 99, 100, 119
Cox’s proportional hazard model. See proportional hazards model
cross-validated Kaplan-Meier curves, 19f, 121–123, 127
cross-validation, 68, 109–111
adaptive signature design, 75–79
error rate, 110–111
leave-one-out-cross-validation, 109–111
log-rank statistic, 128
10-fold cross-validation, 78
data safety monitoring committee, 5
diagonal linear discriminant analysis, 115–116
enrichment design, 35–43
sample size planning, 42–43
standard design versus, 36–40
test performance/specificity, influence on, 40
trastuzumab study, 39
fall-back analysis, 47–48
false negatives
in enrichment designs, 42
in intention to treat analysis, 2
in optimal two-stage design, 26
false positives
in gene detection, 12–13
in intention to treat analysis, 2
in optimal two-stage design, 26
feature selection
class prediction, 108, 113–114
univariate gene selection, 79
Fisher linear discriminant analysis, 115
futility analysis, interim, 50, 56–57
gene finding, for prognostic classifier, 12–13
Gene Expression Omnibus, 78
gene expression profiles, to develop/validate prognostic classifiers, 105–128
gene shaving, 114
genetic algorithms, 114

genomics, vii, 46, 89

goodness of fit, 13–14

hazard function. See proportional hazards model

indication classifier, 46–47

intention to treat, 2

interaction design, 48

interaction tests, 48–49, 54

interim futility analysis, 50, 55–57

intermediate endpoint, 1, 25, 57

k-nearest neighbor classification, 117

Kaplan-Meier survival curves, 14–15, 16f, 18, 96, 121, 125

cross-validated Kaplan-Meier curves, 19f, 121–123, 127

KRAS mutation status, 35, 83, 88

L1 penalized proportional hazards regression, 127

labels, class, 107–108, 109, 110, 111

leave-one-out-cross-validation (LOOCV), 109–111

likelihood

full likelihood, 99

maximum, 98, 99

partial, 100

linear discriminant analysis, 115–116

linear regression, 97, 98

LOE (Level of Evidence) Scale, 83

log-rank test, 76–79

cross-validated statistic, 128

logistic regression, 28–29, 99

MammaPrint score, 11, 22, 33

Mann Whitney test, 113

marker strategy design, 21f, 20–21

matrix, covariance, 71, 99, 100, 119

maximum likelihood, 98, 99

medical utility, 11, 20, 85

microarray analysis, 78. See also prognostic classifiers, based on high dimensional data

MINDACT clinical trial, 22–23

misclassification rate, 116, 118

modified marker strategy design, 22f, 21–22, 23

molecularly targeted drugs, 1

multicenter clinical trials, 5

nearest centroid classification, 117

nearest neighbor classification, 116, 117–118, 125

nearest shrunk centroid classification, 118–119

noise variables, 108

non-informative prior distributions, 103

nonparametric tests, 93

Oncotype DX recurrence score, 11, 22–23, 35

one-sided p-value, 3

optimal two-stage phase II design, 32–33

over-fitting, 68

p-value, 3

one-sided, 3

two-sided, 3

partial least squares, 114, 120–121

partial likelihood, 100

penalized logistic regression models, 72, 119

permutation test, 65–67, 91–93
<table>
<thead>
<tr>
<th>Page 142</th>
<th>Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>phase 2 trials, 25–33</td>
<td>single candidate, 26–28</td>
</tr>
<tr>
<td>Battle I trial in advanced non-small cell lung cancer, 30–33</td>
<td>and study size, 4</td>
</tr>
<tr>
<td>Bayesian adaptive designs for randomized phase 2 trials, 30–33</td>
<td>predictive classifiers</td>
</tr>
<tr>
<td>endpoint as progression-free survival, 30</td>
<td>defining, 35</td>
</tr>
<tr>
<td>logistic regression analysis, multiple candidate biomarkers, 28–29</td>
<td>development of, 106–107</td>
</tr>
<tr>
<td>predictive biomarker design, single candidate binary, 26–28</td>
<td>fall-back analysis and, 47–48</td>
</tr>
<tr>
<td>predictive biomarkers design, one or more binary candidates, 26</td>
<td>indication classifier, 46–48</td>
</tr>
<tr>
<td>purpose of, 1, 25–26</td>
<td>randomized trial comparing new drug to control regimen, 36</td>
</tr>
<tr>
<td>two-stage design, 26</td>
<td>test performance/specificity, influence on enrichment design, 40</td>
</tr>
<tr>
<td>phase 3 trials</td>
<td>predictive pre-specified binary classifier,</td>
</tr>
<tr>
<td>endpoint, 2</td>
<td>test positive/test negative patients, 45–57</td>
</tr>
<tr>
<td>intention to treat principle, 2</td>
<td>adaptively modifying types of patients accrued, 55–57</td>
</tr>
<tr>
<td>interim analyses, 4–5</td>
<td>interaction tests, 48–49</td>
</tr>
<tr>
<td>overview of, 1–5</td>
<td>probabilistic indication classifier, 49–52</td>
</tr>
<tr>
<td>pivotal, 36</td>
<td>sample size planning, 52–55</td>
</tr>
<tr>
<td>power, 3–4</td>
<td>probabilistic indication classifier, 49–52</td>
</tr>
<tr>
<td>purpose of, 45</td>
<td>adaptive threshold design and, 60</td>
</tr>
<tr>
<td>sample size, 3–4</td>
<td>evaluation of sensitivity/specificity of classifier, 52</td>
</tr>
<tr>
<td>statistical significance of, 3</td>
<td>probability of early termination (PET), 32t, 33t</td>
</tr>
<tr>
<td>subset analyses, 5</td>
<td>probabilistic class prediction, 118–120</td>
</tr>
<tr>
<td>population sampling model, 93–94</td>
<td>refinement, 119–120</td>
</tr>
<tr>
<td>pre-specified algorithmic analysis plan, 72</td>
<td>prognostic biomarkers</td>
</tr>
<tr>
<td>predictive biomarkers</td>
<td>classification error, 13</td>
</tr>
<tr>
<td>designs based on single candidate biomarkers, 65–68</td>
<td>false discoveries, 12–13</td>
</tr>
<tr>
<td>designs for development/validation of multivariate classifiers, 68–79</td>
<td>goodness of fit and, 13–14</td>
</tr>
<tr>
<td>identification and validation of, 35–36</td>
<td>medical utility of, 11</td>
</tr>
<tr>
<td>logistic regression model, 28–29</td>
<td>sample size, 13</td>
</tr>
<tr>
<td>molecularly targeted therapy, 25, 53 multiple, 65–79</td>
<td>split-sample approach to avoid bias, 14–15. See also prognostic classifiers</td>
</tr>
<tr>
<td>one or more binary candidates, 26</td>
<td>prognostic classifiers</td>
</tr>
<tr>
<td>predictive classifiers, 68</td>
<td>based on high dimensional data, 105–128</td>
</tr>
</tbody>
</table>
combined models, 127–128
evaluating whether, improves on existing prognostic factors, 126–128
marker strategy design, 20–21, 21f
medical utility of, 20
modified marker strategy design, 22f, 21–22, 23
univariate gene selection, 79
validation studies of, 20–23
prognostic factor studies, 11
proportional hazards model, 96–97, 99–101, 120
baseline cumulative hazard, 122–123
baseline hazard function, 30, 120
endpoint as progression-free survival, 30
hazard function, 96, 100
hazard ratio, 15
L1 penalized proportional hazards regression, 127
sample size planning, 4
prospective-retrospective design, 83–85
randomization
adaptive, 31
Bayesian designs and, 30–33, 49–50
stratified, 30, 45
re-sampling, 14, 125
re-substitution estimate, 111–112
of error rate, 111–112
regression modeling, 11, 97
linear regression, 97, 98
logistic regression, 28–29, 99
proportional hazards regression, 99–101
right-censored data. See survival data
ROC (receiver operating characteristic) curve, 124–125
cross-validated ROC, 125
sample size planning, 12–13, 52–55
and adaptive randomization, 31
adaptive threshold design, 62–63
enrichment design, 42–43
optimal two-stage design, 32–33
proportional hazards model, 4
sample splitting, 68
shrunken centroid classification, 117
split-sample method, 109
statistical power, 3–4, 96
statistical significance, 91–94
one-sided p-value, 3
permutation significance test, 65–67
threshold significance level, 4–5, 12, 113
two-sided p-value, 3
stratification design, 45
strong null hypothesis, 59–60, 76
study-wise type I error, 48–49
supervised principle component classifier, 114
support vector machines (SVMs), 116
survival analysis, 96–97
survival risk prediction, 120–125
cross-validated Kaplan-Meier curves, 19f, 121–123, 127
Kaplan-Meier survival curves, 14–15, 16f, 18, 96
survival risk classifiers, 120–121
time-dependent receiver operating characteristic curves, 124–125
t-test, 113
TAILORx clinical trial, 32–33, 35, 86
threshold significance level, 4–5, 12, 113
time-dependent receiver operating characteristic curves, 124–125
time-to-event endpoint, 54–55, 76–79
tuning parameters, 108
two-way analysis of variance, 48
type I error, 5, 83
  in adaptive signature design, 73
  and Bayesian methods, 103
  defined, 4–5
  study-wise, 48–49
  two-sided, 12
  without specifying cut-point in advance, 60

validity
  analytical, 11, 88
  calibration as measure of, 119–120
  clinical, 11, 85
  medical utility, 11, 20, 85
  of regression model, 101
  validation studies of prognostic classifiers, 20–23