Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Med Decis Making. Author manuscript; available in PMC 2017 July 01.

Published in final edited form as:

Med Decis Making. 2016 July ; 36(5): 594–603. doi:10.1177/0272989X15601998.

Comparative Effectiveness of Biomarkers to Target Cancer Treatment: Modeling Implications for Survival and Costs

Jeanette K. Birnbaum[1], **Foluso O. Ademuyiwa**[2], **Josh J. Carlson**[3,4], **Leslie Mallinger**[5], **Mark W. Mason**[4], and **Ruth Etzioni, PhD**[1,5,6]

[1]Department of Health Services, University of Washington, Seattle WA

[2]Department of Medicine, Washington University in St Louis, St Louis, MO

[3]Pharmaceutical Outcomes Research and Policy Program, University of Washington, Seattle WA

[4]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle WA

[5]Department of Health Administration, University of Washington, Seattle, WA

[6]Department of Biostatistics, University of Washington, Seattle WA

Abstract

Background—Biomarkers used at the time of diagnosis to tailor treatment decisions may diffuse into clinical practice before data become available on whether biomarker testing reduces cancer mortality. In the interim, quantitative estimates of the mortality impact of testing are needed to assess the value of these diagnostic biomarkers. These estimates are typically generated by customized models that are resource-intensive to build and apply.

Methods—We developed a user-friendly system of models for Cancer Translation of Comparative Effectiveness Research (CANTRANce) to model the mortality impact of cancer interventions. The Diagnostic Biomarker module of this system projects the mortality impact of testing for a diagnostic biomarker, given data on how testing impacts treatment recommendations. Costs and quality of life outcomes may also be modeled. We applied the Diagnostic Biomarker module to two case studies to demonstrate its capabilities.

Results—The user interface (http://www.fhcrc.org/cantrance) allows comparative effectiveness researchers to use the Diagnostic Biomarker module of CANTRANce. Our case studies indicate that the model produces estimates on par with those generated by customized models and is a strong tool for quickly generating novel projections.

Limitations—The simple structure that makes CANTRANce user-friendly also constrains the complexity with which cancer progression can be modeled. The quality of the results rests on the quality of the input data, which may pertain to small or dissimilar populations or suffer from informative censoring.

Reprint requests to: Jeanette Birnbaum, Department of Health Services, University of Washington, Box 357660, Seattle WA 98195-7660. Tel: +1.206.667.3191. Fax: +1.206.667.7264. ; Email: kurian@uw.edu

Conflicts of interest: None.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Conclusions—The Diagnostic Biomarker module of CANTRANce is a novel public resource that can provide timely insights into the expected mortality impact of testing for diagnostic biomarkers. The model projections should be useful for understanding the long-term potential of emerging diagnostic biomarkers.

Introduction

Over the last decade, advances in genomic research have produced biomarkers to target cancer treatments to a patient's unique molecular characteristics.[1–4] Testing patients for biomarkers that either predict response to particular cancer therapies or supply prognostic information can support tailored treatment decisions that improve outcomes and/or reduce toxicity. Established examples of such biomarkers include the Her-2/Neu oncogene that predicts positive response to trastuzumab for breast cancer and the KRAS mutations that indicate resistance to anti-EGFR treatments for metastatic colorectal cancer.[5] In both of these examples, randomized controlled trials (RCTs) of patients who are positive for the biomarkers in question have demonstrated better survival with tailored treatment.[4,6–10]

In contrast, emerging biomarkers may diffuse into clinical practice long before survival data become available. Depending on the cancer, more than a decade may pass between early-stage studies of a diagnostic biomarker and an RCT that confirms benefit of testing and tailored treatment on survival, if such a study is even done.[11] In the interim, comparative effectiveness (CE) studies typically collect data on how testing for a biomarker impacts an intermediate endpoint such as treatment recommendations.[12–17] While data on treatment recommendations can suggest that testing has affected treatment, these data do not address the ultimate question of whether testing reduces cancer mortality.

Quantitative estimates of the mortality impact of testing are needed to assess the value of these biomarkers, which we will call "diagnostic biomarkers" given their use at the time of diagnosis to inform treatment choices. One way to meet this need is to conduct a complementary modeling study that translates changes in treatment recommendation into projected effects on mortality. This strategy of using models for comparative effectiveness research (CER) when observational data are unavailable has been highlighted as both important and underutilized in cancer research.[18,19] Moreover, the fast pace of biomarker development has created a demand for CER that cannot be fully met by observational studies: when competing diagnostic biomarkers emerge, there may be more head-to-head comparisons than can be feasibly investigated. This environment is a perfect opportunity for modeling to supply otherwise unattainable comparative effectiveness results.

A practical obstacle to modeling the impact of diagnostic biomarkers is that the investigators who conduct CE studies of how diagnostic biomarkers impact treatment typically do not have the additional resources to conduct the complementary modeling study translating their results to mortality. On the other hand, research groups with modeling expertise often conduct modeling studies using aggregate data from the literature because they are not connected with investigators who have primary data. In addition, many modeling studies are highly application-specific and customized to the particular question at hand, preventing them from being quickly re-applied as new diagnostic biomarkers emerge.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

To address this gap, we have developed a system of statistical simulation models for Cancer Translation of Comparative Effectiveness Research (CANTRANce) that facilitates the modeling of the mortality impact of interventions, given CE studies of typical intermediate endpoints, i.e. the common pre-mortality outcomes used to demonstrate the effectiveness of interventions. CANTRANce can model the mortality impact of several intervention-intermediate endpoint pairs: preventive interventions and disease incidence; screening tests and sensitivity and specificity of detection or stage distribution of detected cancers; diagnostic biomarkers and treatment distributions; and treatments and disease recurrence. Users enter data from a CE study of how an intervention impacts an intermediate endpoint, and the modeling system prompts them for additional data from the literature or other studies. The system then extrapolates from the intermediate endpoint to the projected impact of the intervention on mortality. In this paper, we introduce the Diagnostic Biomarker module of CANTRANce and use it to evaluate a diagnostic biomarker for breast cancer. We show how CANTRANce can flexibly model the impact of a new diagnostic biomarker on mortality, as well as on quality-adjusted life-years and costs.

Methods

CANTRANce Framework

In all CANTRANce models, we use an individual-level microsimulation approach in which individuals transition through health states, and transitions occur in continuous time rather than at specified time intervals (Appendix A: Glossary of Terminology).[20,21] The models simulate a virtual population and project outcomes for two scenarios, one in which individuals do receive the intervention and one in which they do not. The impact of the intervention on the intermediate endpoint is based on a CE study ending at that endpoint. We model each person's subsequent series of health states or "life history" by projecting the dates of key events such as disease progression and mortality.[20,21]

The Diagnostic Biomarker (Diagnostics) module of CANTRANce evaluates the impact of testing for a new diagnostic biomarker to target treatment (Figure 1). The model compares two scenarios: 1) all individuals receive a one-time testing intervention using the diagnostic biomarker and 2) no individuals receive testing. The foundation for the model is a CE study of how the testing intervention impacts the intermediate endpoint, treatment recommendations. Based on this CE study, the testing intervention and no testing scenarios each have different distributions of two possible treatments: a standard treatment, and a tailored treatment whose indication changes after test results are known. In order to project to mortality, the model prompts the user for the necessary additional information on expected cancer mortality following each treatment option. Alternatively, the user may project to mortality through a second intermediate endpoint of disease recurrence.

The precise inputs and steps involved in the projection depend on certain choices available to the user. These elements are described in detail below from the perspective of a user running the model via the user-friendly Windows application interface, downloadable from http://www.fhcrc.org/cantrance. The interface prompts the user for the necessary inputs and then runs the model locally using an open-source R package, *cantrance* (v 1.1). The results and open-source code for the model are stored on the user's local machine.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

General modeling selections—The user must specify the number of simulations and years at which to evaluate survival. Each simulation represents one study linking testing to treatment and mortality, and multiple simulations capture the stochastic uncertainty due to the simulation process.[22] Users can also define multiple follow-up time points for the projection, e.g. 10-year as well as lifetime follow-up.

Constructing the virtual population—The Diagnostics model begins by generating a virtual population that mimics the population in the CE study. At minimum, each individual in the virtual population needs to be assigned a gender, age at treatment, and either 1) treatment recommended under the testing intervention versus under no testing, for paired CE studies in which treatment was assessed for the same patients under both scenarios, or 2) treatment recommended and an intervention indicator, for unpaired CE studies in which the testing intervention versus no testing was studied using different patients. Other relevant covariates may also be included. The two options for generating this population are:

1. *Specify covariate summaries*. If the user only has population-level summary statistics for covariates reported in the CE study publication, the Diagnostics model can use those to generate a virtual study population for each simulation in which individuals have covariate values that are consistent with the published data at the population level. Both marginal and joint distributions of categorical and continuous covariates can be accommodated. The Diagnostics model will assume independence between covariates without specified joint distributions.
2. *Bootstrap from individual-level data*. If individual-level data from the CE study are available to the user, the entire CE study population can be entered into the model. For each simulation, the model will bootstrap from this population to define the virtual study population. Bootstrapping preserves all the interrelations between covariates in the original study. While the default bootstrap is unweighted, the user may alternatively specify weights for the resampling in order to model the impact of testing on a population with a particular covariate distribution.

Modeling treatment received—The virtual population includes an indicator of treatment recommended under each scenario, and this value can be used directly. Alternatively, if individual-level data are available, the user may choose to run a logistic regression of treatment received, separately for each scenario, based on available predictors in the data. For the testing intervention scenario, the test result is a required predictor. In each simulation, CANTRANce fits the regression and uses the results to simulate treatment recommended under each scenario. If the CE study is very small, the regression-based treatment distributions for the virtual population may be more stable than the observed values.

Modeling time from treatment to mortality—This step projects the dates of cancer and other-cause mortality for each virtual individual. The cause of death is determined by the event that occurs first.

The Diagnostics model projects times to event using a single, user-specified survival statistic, e.g., an event rate or k-year survival. This statistic is translated into an exponential

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

survival curve from which the model simulates times to event for each individual. Hazard ratios may be specified to modify this curve for particular covariate values.

It is important to note that the survival statistic is interpreted as a *net* statistic, that is, in the absence of other-cause death (Appendix A: Glossary of Terminology). Users may find some net statistics reported directly in the literature. In particular, relative survival is an appropriate net statistic[23] routinely reported by SEER Cancer Statistics Review.[24] The Surveillance, Epidemiology and End Results (SEER) database[25] provides another venue for deriving net survival statistics. Alternatively, a composite statistic such as overall survival or disease-free survival can be used to approximate a net statistic when the risk of death from other causes is relatively low in the study cohort (e.g., in a younger cohort).

The user may choose to use disease recurrence as an intermediate endpoint within the time to mortality projection. If so, the Diagnostics model requires one survival statistic for the time from treatment to recurrence, and another statistic for time from recurrence to mortality. Otherwise, the user specifies one survival statistic describing time from treatment to mortality. In either case, the user must specify at least one hazard ratio describing the benefit of tailored versus standard treatment on the event (recurrence and/or mortality). Additional hazard ratios may be specified if survival or the benefit of tailored treatment varies by other covariates.

Age at other-cause death is simulated from US cohort life tables.[26] A hazard ratio on the life table may be specified if the CE study population is thought to have all-cause survival that differs from the general population.[27]

Modeling quality of life and cost—Because some diagnostic biomarkers primarily function to spare patients unnecessary treatment and thus improve quality of life, the Diagnostics model also allows users to specify annual costs and utility weights for key periods during the life history: the first year of treatment, subsequent years until mortality, and the year of death. If recurrence is used in modeling survival, cost and utilities can be further partitioned by the onset of recurrence. All costs and utility weights can vary by tailored versus standard treatment. The user may also specify a discount rate to discount future gains to their present value.[28]

Quantifying uncertainty—Each run of the Diagnostics model performs multiple simulations using the same parameter set.[20] Stochastic (i.e., Monte Carlo) uncertainty is quantified by summarizing results across simulations using the mean as the point estimate and the 2.5% and 97.5% quantiles as the 95% uncertainty interval. Sensitivity to parameter values can be easily investigated by changing parameters through the user interface. Formal one-way and probabilistic sensitivity analyses that systematically investigate parameter sets are also possible outside of the interface, using some customized R programming. For sensitivity analyses, we again describe uncertainty using the 2.5% and 97.5% quantiles across runs to define the 95% uncertainty interval.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Case Studies

The Diagnostics model is designed to be a high-quality, flexible tool to project long-term outcomes following testing with a diagnostic biomarker. We developed two case studies that highlight these features. The first case compares the Diagnostics model's performance to that of a previously developed customized model, and the second case provides novel CE projections in a setting where there is no existing model.

Both cases are founded on CE studies of how a testing intervention using a 21-gene recurrence score (RS) for breast cancer impacts treatment recommendations for adjuvant chemotherapy plus hormone therapy (tailored treatment) versus hormone therapy alone (standard treatment). The RS classifies patients as being at low, intermediate or high risk of disease recurrence. The CE studies describe how physician treatment recommendations change with and without RS knowledge. The Diagnostics model translates these changes in treatment to their impact on mortality, quality-adjusted life years and costs. In both cases, we discount outcomes using the standard rate of 3%.[29]

Case 1: Replication of a customized model—In Case 1 we replicated projections made by a customized model developed by Reed et al to compare using the RS to guide treatment (the testing intervention, or "with RS") to NCCN-guideline-based treatment (no testing, or "without RS").[30] Reed et al used a Markov cohort decision model to project outcomes for these two scenarios based on a published CE study in which medical oncologists made two treatment recommendations for node-negative, ER-positive patients, first without the patients' RS and then with the RS.[12,31] Because the model used published CE data, no individual-level data were available and we thus used the covariate summaries option of the Diagnostics model to replicate the CE study population reported in Reed et al. We projected time to mortality, total costs from the societal perspective, and quality-adjusted life-years using recurrence as a second intermediate endpoint (Table 1). We additionally replicated the probabilistic sensitivity analysis (PSA) using the same standard errors and distributional assumptions (Table 1). As a sensitivity analysis, we investigated the impact of customizing the Diagnostics model to match the Reed model on two final elements: their period life table (versus our cohort life table) and their assumption that recurrence is possible only in the first 10 years (versus our unlimited possibility of recurrence).

Case 2: Novel projection of the impact of diagnostic testing—In Case 2, we did a novel projection of the impact of the RS using individual-level data from a CE study by Ademuyiwa et al of how testing impacted treatment recommendations in node-negative, ER-positive women.[14] The CE study compared treatment recommendations made with RS knowledge (the testing intervention, or "with RS") to those made in a retrospective chart review in which the RS was withheld from the medical oncologist (no testing, or "without RS"). In this case, the individual-level CE data were made available to us, so we used the individual-level CE data option to construct the virtual population and projected mortality, total costs and quality-adjusted life-years using the same parameters as in Case 1 (Table 1). We also performed a PSA using the same standard errors and distributional assumptions as in Case 1.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Tutorial

To facilitate use of the Diagnostics model, we provide a tutorial in the Appendix that shows how to replicate the Case 1 and Case 2 models through the user interface and provides additional technical details and modeling guidance. The tutorial also demonstrates how advanced users can access the R code. This written tutorial is accompanied by a video tutorial to familiarize users with the interface (http://www.fhcrc.org/cantrance). Due to data privacy restrictions, the Case 2 replication in the tutorial uses a public-access approximation of the true individual-level data, generated from the information published in Ademuyiwa et al.[14]

Results

User Interface

The interface for Diagnostics model walks the user through each of the modeling steps outlined in the *CANTRANce Framework* section of the Methods (Figure 2). The user specifies modeling choices in a scrolling left panel through a combination of selection boxes, user-defined tables, and numeric entry boxes. Individual-level CE data can be entered using a comma-separated text (.csv) file. Once the CE study data are entered, the left panel updates to reflect those data. The right panel of the interface allows the user to navigate between sections of the model, view help files, and run the model with a viewport to the R session.

Results of Case Studies

The CE studies used as inputs to the Diagnostics model for each case had different population ages and risk distributions (Table 2). All Case 1 patients were assumed to be 55, following the Reed model.[30] Average age in Case 2 was also 55, but ages ranged from 29 to 82. About half of patients were classified as low RS in both cases, but the Case 2 data had very few high RS patients compared to Case 1.

Use of the RS decreased adjuvant chemotherapy among low RS women and increased it among high RS women in both samples. Case 1 data additionally indicated an increase in adjuvant therapy among intermediate RS women with use of the RS (Table 2).

The models for both CE studies projected that testing with the RS increases life-years and quality-adjusted life-years (QALYs) at reasonable cost, but with substantial uncertainty (Table 3). Benefits of using the RS were greater for the Case 1 models due to the greater proportion of high-risk patients in the population, for whom use of chemotherapy increased from 67% to 100% in Case 1 versus the smaller increase of 88% to 96% in Case 2 (Table 2). In addition, more intermediate RS women were switched onto chemotherapy with the RS in Case 1 (Table 2) and this was modeled as beneficial for all of them (Table 1). In both the high- and intermediate-RS groups, the survival benefits of chemotherapy outweighed its negative impact on quality of life to result in QALYs gained on average. However, the additional courses of chemotherapy also resulted in greater incremental costs in Case 1. In Case 2, the primary effect of testing was to decrease chemotherapy in low RS women (Table

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

2). Costs saved by avoiding unnecessary chemotherapy in Case 2 almost completely offset the cost of testing, on average (Table 3).

The Case 1 results projected greater incremental life-years and QALYs than those reported by Reed, but in our sensitivity analysis we were able to closely replicate the Reed results (Table 3). When we matched the Reed model's life table and assumption limiting recurrence to the first 10 years, we projected 0.23 incremental life-years, 0.18 incremental QALYs, $1,920 incremental total costs per person for with RS versus without RS, and an incremental cost-effectiveness ratio (ICER) per QALY of $10,772. Reed projected 0.19 incremental life-years, 0.16 incremental QALYs, $1,741 incremental total costs, and an ICER per QALY of $10,881.[30] The remaining differences between our Case 1 sensitivity analysis and the Reed results are due to the differences in model structures. Our microsimulation model uses a continuous-time approach and represents a population in which parameter inputs define distributions (e.g., exponentially-distributed survival). Their Markov model uses 6-month cycles and represents a homogeneous cohort subject to the exact parameter inputs.

Discussion

The Diagnostic Biomarker module of CANTRANce is a powerful new tool for understanding the likely impact of new diagnostic biomarkers to target cancer treatment. As a free, open-source modeling platform, it is a unique public resource to support timely comparative effectiveness research on emerging diagnostic biomarkers. The user-friendly interface allows investigators to extend their CE results of how testing impacts treatment to mortality, quality-adjusted mortality and costs without needing to commission a customized model. This is made possible by the variety of modeling options available to the user within the Diagnostics model. The model can flexibly accommodate CE data in multiple forms as well as conform to various other user needs, such as the use of disease recurrence as a second intermediate endpoint. The user can also use the interface to easily investigate how altering model assumptions changes results. The availability of all these features in one model holds great potential for quick projections as CE studies of how emerging diagnostic biomarkers impact treatment become available. Head-to-head comparisons of competing diagnostic biomarkers may reveal expected differences—or a lack thereof—that can be used to prioritize further research.

The two case studies presented in this paper validate the capability of the Diagnostics model to produce projections on par with customized models and provide a novel projection of the impact of testing with a diagnostic biomarker. The Case 1 model closely replicated the Reed model once all assumptions were matched. The Case 2 model suggested less benefit but also lower costs with testing than Case 1. The different results for the two cases highlight that the projections can be sensitive to the CE study on which they are based. Case 1 projected greater benefit and costs than Case 2 due to the different population risk distributions and the differences in the impact of testing indicated by the two CE studies. The interpretation and application of these projected outcomes must thus be considered in context of the original CE study population.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The microsimulation approach we chose for CANTRANce allows us to model the CE study population with realistic variation. In a Markov model like the Reed model, population characteristics are typically simplified in order to maintain a reasonable number of nodes and branches. Microsimulation allows representation of the variation in characteristics across individuals in a population. For example, in a Markov model, if the probability of being recurrence-free at 10 years is 0.968, each simulated individual is subject to that exact probability of being recurrence-free. In a microsimulation model, 0.968 is used to define a distribution for recurrence-free survival, and individuals are assigned different times to recurrence that, when summarized at the population level, equate to a probability of 0.968 at 10 years. One consequence of modeling population variation in this manner is greater uncertainty in results, as seen in Case 1 uncertainty intervals compared to those reported in Reed et al.[30]

The CANTRANce approach of projecting mortality based on CE studies of intermediate endpoints also has its limitations. As with all models, the quality of the projection will reflect the quality of the input data and the accuracy of the assumptions involved. Study populations, such as the ones used in the two case studies, may be small and not representative of major populations of interest. Survival data used to inform model parameters may not be available from populations that resemble the study population. Net survival in particular is not always reported in publications and often must be approximated. In both case studies, we used recurrence-free survival from published studies to approximate the net survival distribution for time from diagnosis to recurrence. This may overestimate the incidence of recurrence. We used relative survival for advanced SEER cases to project the net survival following disease recurrence. So long as survival for recurrent cases is reasonably approximated by survival for newly-diagnosed advanced cases, this approximation should be reasonable. Even when available, net disease-specific survival may suffer from informative censoring due to competing events such as other-cause death. Finally, the modeling approach of projecting mortality from intermediate endpoint(s) requires that those endpoints behave as surrogates for mortality. We recognize that endpoints for different cancers may vary in the degree to which they are well-established as surrogates.[32] In our examples, we used disease recurrence as a surrogate for breast cancer mortality, a reasonable approach given the large contribution of recurrent cases to mortality and the low likelihood of cure after recurrence.[33] The model also allows for the inclusion of additional covariates for projecting mortality after the intermediate endpoint. Furthermore, if users extend the model to include quality-of-life and cost data, the quality of those inputs will also impact the accuracy of the results.

CANTRANce also has limitations imposed by its generic form. The simple model structure that makes CANTRANce user-friendly also limits its ability to incorporate more complex features represented in customized models. A primary simplification is the shape of the time to event curve, which is constrained to follow an exponential distribution in CANTRANce. While mortality is typically better characterized by more flexible survival curves, specifying these curves would require additional data that may be unavailable to the user. Moreover, any parametric survival curve is likely a simplification of the mortality process. In the case of ER positive breast cancers, a recent study suggests that the constant mortality hazard implied by the exponential distribution is likely a reasonable approximation to observed

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

survival for these cancers.[34] Other simplifications include the number of key events represented in the disease process and lack of explicit representation of adverse events or heterogeneity in within treatment groups. In the Diagnostics model, we also do not distinguish between treatment recommendations and final treatment decisions, assuming that the recommendations are generally followed. What is gained by losing more complex features is the ability to quickly re-use the model for similar applications without having to build the model from scratch.

The Diagnostics model is a platform for future studies of the value of diagnostic biomarkers for targeting cancer treatment. New diagnostic biomarkers will continue to diffuse into clinical practice before their impact on mortality can be observed. In CANTRANce, we offer a blueprint and modeling system for the extrapolation process. We hope that this will empower CE investigators to project the consequences of their interventions beyond intermediate endpoints to long-term endpoints that can inform the adoption of diagnostic biomarkers.

Acknowledgments

Financial support: Financial support for this study was provided by grant RC4 CA155806 from the National Cancer Institute and grants 1K12CA167540 through the National Cancer Institute and UL1 TR000448 through the Clinical and Translational Science Award program of the National Center for Advancing Translational Sciences at the NIH (FOA). The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institute of Health.

Appendix A: Glossary of Terminology

This glossary includes definitions for various modeling-related terms as they apply in the CANTRANce context. References refer to the reference list following the glossary table.

Term	Definition	Reference
Bootstrapping	Sampling individuals with replacement from a dataset to generate a new version of the data that has a different mixture of individuals. Using a different bootstrap for each simulation allows CANTRANce to capture the sampling uncertainty of an input dataset.	Chernick 2003
Crude survival	Survival pertaining to one cause of death only, e.g. breast cancer, accounting for the possibility of dying from other causes. Also known as 1-cumulative incidence of mortality.	Kim 2007
Disease-free survival	Time from study start to the first of multiple possible events: recurrence, relapse, second cancer, or death. Synonym for **recurrence-free survival**.	Gill 2006
Disease-specific survival	Survival pertaining to one cause of death only, e.g. breast cancer. See **net survival** and **crude survival**.	Danieli 2012
Exponential survival curve	A type of **survival curve** that assumes that the rate of events is constant. Can be defined using one parameter only: either a rate, proportion surviving at *k* years after time=0, median survival, or mean survival. These statistics can be converted into one another, e.g. a particular rate implies a particular median survival.	Lee 1997
Hazard ratio	The ratio of the instantaneous rates of two events. Hazard ratio (HR) = 1 indicates identical rates. A hazard ratio for treatment	Lee 1997

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Term	Definition	Reference
	benefit would be a number less than 1 that indicates the degree to which treatment decreases the event (e.g., recurrence) rate.	
Kaplan-Meier method	Method for estimating **net survival**. Involves censoring deaths from other causes.	Danieli 2012
Life history	The life events modeled for each virtual individual in a microsimulation model, e.g. dates of disease progression, death from cancer, and death from other causes.	Rutter 2011
Life table (period or cohort)	Data indicating life-expectancies at regular age intervals between birth and death. Period life tables are constructed empirically using age-specific death rates in a given year and represent a synthetic cohort that experiences death rates from one snapshot in time over their lifetime. In contrast, cohort life tables reflect the change in age-specific death rates over calendar time but require projection of death rates for future years when mortality is not yet observable.	Guillot 2011
Median survival	Time after the index time=0 on a survival curve at which 50% of the population has experienced the event	Lee 1997
Microsimulation modeling	Individual- or patient-level model in which life events or a "**life history**" are modeled in continuous time rather than discrete time intervals. Generating these times involves a random process for which uncertainty can be quantified by summarizing results over multiple **simulations**.	Rutter 2011
Monte Carlo uncertainty	Uncertainty due to random-number generation in a **simulation** run. Summarizing population results across simulations allows quantification of this mode of uncertainty.	Rutter 2011, Briggs 2012
Net survival	Survival pertaining to one cause of death only, e.g. breast cancer, as if it were the only possible cause of death (i.e. In the absence of other causes of death).	Danieli 2012
Overall survival	Survival pertaining to all causes of death.	
Recurrence-free survival	Synonym for **disease-free survival.**	Gill 2006
Relative survival	A type of net survival. Involves comparing the **overall survival** of a cancer cohort to overall survival in a general population cohort.	Danieli 2012
Simulation	One instance or "run" of a microsimulation model. Each run involves a random process to generate events in the life history. Individuals in the model thus receive different event times in different simulations. However, in each simulation, the distribution of each event time across individuals in the population remains consistent with the population parameter used as the model input (e.g. a mortality rate).	Rutter 2011
Simulation uncertainty	Synonym for **Monte Carlo uncertainty**.	Rutter 2011, Briggs 2012
Stochastic uncertainty	Synonym for **Monte Carlo uncertainty**.	Rutter 2011, Briggs 2012
Survival curve	Describes the distribution of times to an event, e.g. recurrence or mortality, in terms of the proportion of a population surviving (*not* experience the event) at a given time point after an index time=0.	Lee 1997

Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD, et al. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6. Med Decis Making. 2012 Oct;32(5):722–32.

Chernick MR, Friis RH. Introductory Biostatistics for the Health Sciences: Modern Applications Including Bootstrap. John Wiley & Sons; 2003. 427 p.

Danieli C, Remontet L, Bossard N, Roche L, Belot A. Estimating net survival: the importance of allowing for informative censoring. Stat Med. 2012 Apr 13;31(8):775–86.

Gill S, Sargent D. End points for adjuvant therapy trials: has the time come to accept disease-free survival as a surrogate end point for overall survival? Oncologist. 2006 Jun;11(6):624–9.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Guillot M. Period Versus Cohort Life Expectancy. In: Rogers RG, Crimmins EM, editors. International Handbook of Adult Mortality [Internet]. Springer Netherlands; 2011 [cited 2015 Feb 6]. p. 533–49. Available from: http://link.springer.com.offcampus.lib.washington.edu/chapter/10.1007/978-90-481-9996-9_25

Kim HT. Cumulative Incidence in Competing Risks Data and Competing Risks Regression Analysis. Clinical Cancer Research. 2007 Jan 15;13(2):559–65.

Lee ET, Go OT. Survival Analysis in Public Health Research. Annual Review of Public Health. 1997;18(1):105–34.

Rutter CM, Zaslavsky AM, Feuer EJ. Dynamic microsimulation models for health outcomes: a review. Med Decis Making. 2011 Feb;31(1):10–8.

Appendix B: Tutorial for Using the Diagnostics Model

Installing the R statistical software and Microsoft Access Runtime

1. Go to http://cran.r-project.org/bin/windows/base and download R for Windows.

2. Open the downloaded .exe file to install R. Use the default installation folder, e.g., C:\Program Files\R\R-3.1.1. Note the location of this folder, as you will need to confirm it later.

3. Go to http://www.microsoft.com/en-us/download/details.aspx?id=39358 to download Microsoft Access Runtime. Follow the “Installation Instructions” section to select the appropriate version.

Note: you will also need an internet connection when you first use the Diagnostics model, to allow the model to auto-download R packages used by the model.

Installing and initializing the Diagnostics model interface, and accessing R code

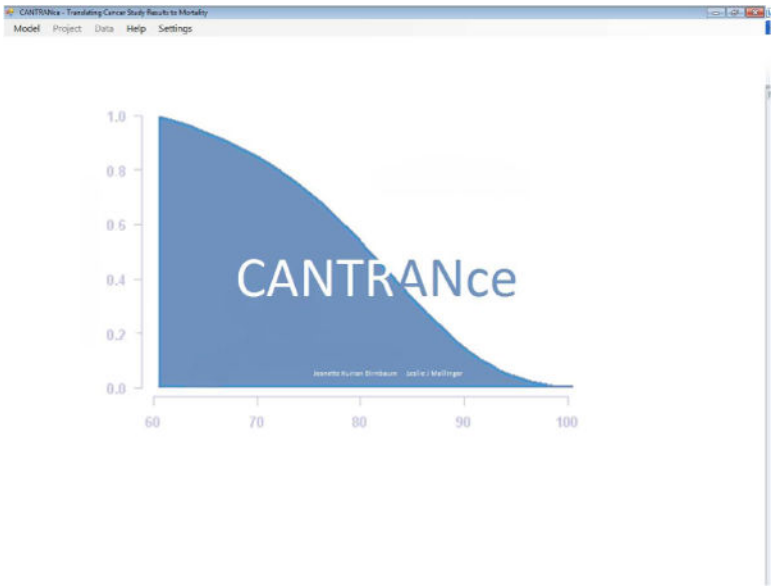
Note: This software was built and tested on 64-bit Windows 7.

1. Go to http://www.fhcrc.org/cantrance and click on the Windows application link. You may register to receive updates if you wish.

2. Watch the video tutorial, linked from the downloads page (13 min). This will familiarize you with how the interface works and preview for you how the Case 1 model can be run using the interface.

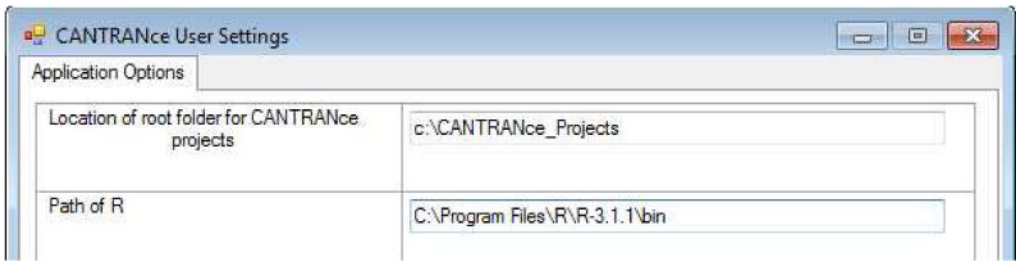
3. Download the setup.exe file.

4. Open the .exe file to install CANTRANce. You will see the welcome screen:



5. Select the Diagnostic Tests model from the Model menu.

6. Go to Settings and edit the “Path of R” to point to the folder where R.exe is stored. If you didn’t choose a special location when installing R, it will most likely be in C:\Program Files \R-3.1.2\bin, as indicated below. You may also edit the “Location of root folder for CANTRANce projects.” Make sure to click “Save” after editing the “Path of R.”



7. Open a New Project from the Project Menu. Within C:\CANTRANce_Projects, create a folder of your choice to store the project. For your first use of the model, we recommend you name your folder “Project1” in order to easily follow later steps in the tutorial.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

8. Use a file browser window to look at the newly created folder. In the Inputs subfolder, you will see an example individual-level CE data file that will be used in the Case 2 projection, as well as a run_file.R that contains the underlying code for the model.

9. To run the Case 1 model, choose “Covariate Proportions” and continue on to the next section, “Running the Case 1 model – Part I.” To run the Case 2 model, choose “Individual-Level Data.” An open-file dialogue box will pop up. Proceed to the section “Running the Case 2 model – Part I” on page 9. If this is your first use of the model, you may want to run the Case 2 model as it requires less setup time.

Note: Once the interface is initialized, in each section, you may hover your mouse over the green **?** to display a detailed description of each parameter in the “Help” box in the lower right panel. The following instructions for running each case will provide background on each parameter. Read the Help for more details.

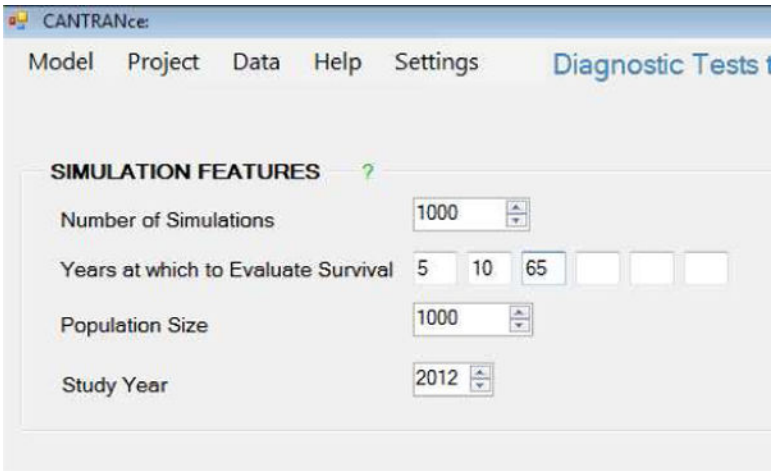
Running the Case 1 model – Part I

Note: Running the Case 1 model involves hand-keying in a large table of CE study data. If you want a shorter introduction to the Diagnostics model, skip to the Case 2 model on page 9.

1. Enter the following parameters in the “Simulation Features” section.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



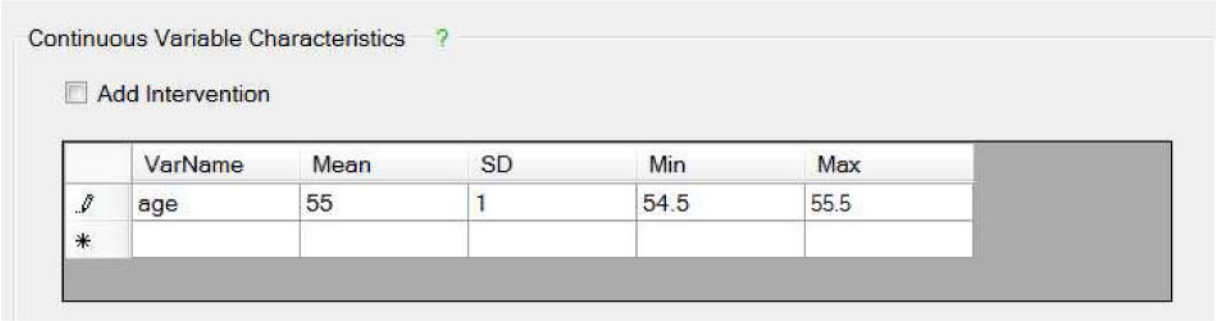
Number of Simulations: 1000 simulations is a standard number of replicates to characterize uncertainty. If you are testing out different models, fewer simulations will produce results faster, but we recommend at least 1000 simulations for a final model. Since the Reed model reported lifetime results, **Years at which to Evaluate Survival** needs a maximum follow-up time that represents lifetime follow-up. Given that all women were aged 55 in the Reed model, we enter 65 years as the maximum follow-up because all women will die by age 120 according to our cohort life table. Because the Reed model presented per-patient rather than population results, the **Population Size** is not very influential and we choose a standard size of 1000. This size impacts the model in two ways: uncertainty due to the simulation process will decrease as (Number of simulations)*(Population Size) increases, and any absolute (versus per-patient) outcomes will refer to this particular population size.[1] We choose 2012 as the **Study Year** given the Reed paper publication date. This parameter will identify birth cohorts for simulated individuals, i.e. women aged 55 in 2012 were born in 1957.

2. Choose "Yes" for **Is this a paired study design?** The Case 1 CE study determined treatment recommendations with and without testing for the *same* patients.

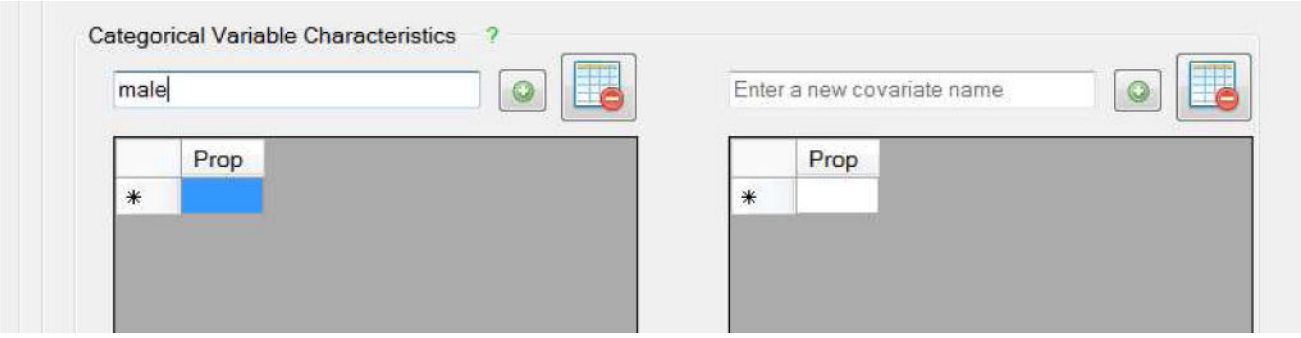
[1]For more on this topic, see Sharif B, Kopec JA, Wong H, Fines P, Sayre EC, Liu RR, Wolfson MC. Uncertainty Analysis in Population-Based Disease Microsimulation Models. Epidemiology Research International. 2012 Jul 25;2012:e610405.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

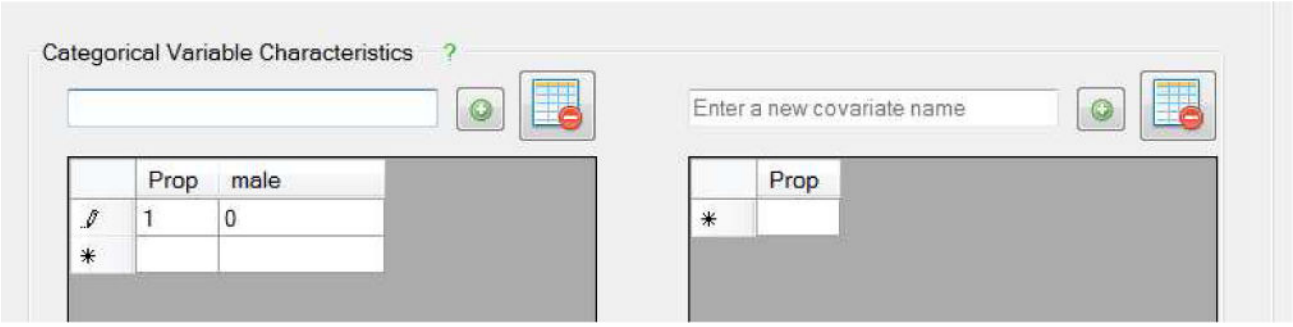
3. Now we will start entering the CE data into the model. Enter “age” as the one continuous variable in **Continuous Variable Characteristics**, as shown below. Since all women were age 55 in the Case 1 model, we will use a truncated normal distribution to normally distribute women’s ages between 54.5 and 55.5 with a mean of 55.0. If the study population had a more complex age distribution, this could be accommodated in several ways – hover over the green ? and read the Help for instructions.



4. Enter “male” as the first categorical variable in **Categorical Variable Characteristics** by typing “male” in as the new covariate name and pressing the green plus button.



Then enter 0 as the value and Prop(portion) as 1. This tells the model that all individuals are male=0, or female. A population with a mixture of sexes could be modeled by adding a male=1 row and specifying the proportions (e.g., a 50-50 split would be Prop=0.5 for both rows).



Author Manuscript Author Manuscript Author Manuscript Author Manuscript

5. Enter the remaining CE study data using the 2nd table entry box in **Categorical Variable Characteristics** using the guidelines below. For complete details on how to enter CE study data using **Categorical Variable Characteristics**, please see the Help for this section by hovering over the green ?. It explains how variables can be simulated using either joint or independent distributions. Here, we will focus specifically on the Case 1 CE data entry, in which we need to specify joint distributions for how chemotherapy was recommended according to patients' risk categories.

We will indicate the receipt of chemotherapy (the tailored treatment) in the with-testing scenario using a variable called **rx_test**, and in the without-testing scenario using a variable called **rx_notest**. Because the Case 1 CE study looked at the impact of testing by NCCN and RS risk groups (see Reed Table 1), we must also specify these risk groups using the variables **nccn_risk** and **rs_risk**.

Enter these four variables (**nccn_risk, rs_risk, rx_notest**, and **rx_test**) as the variable names for the 2nd table. (***Note****: you can resize the column widths to avoid scrolling and use the Tab and Enter keys to navigate within the table. Press the button with the red minus sign to destroy the table and start over.)*

6. Populate this table with 36 rows using the values in the table below. These values are derived from Reed Table 1, by multiplying the proportion of patients in each joint risk group (e.g., NCCN-risk=low and RS-risk=low) by the proportion falling within each of the four possible outcome groups:

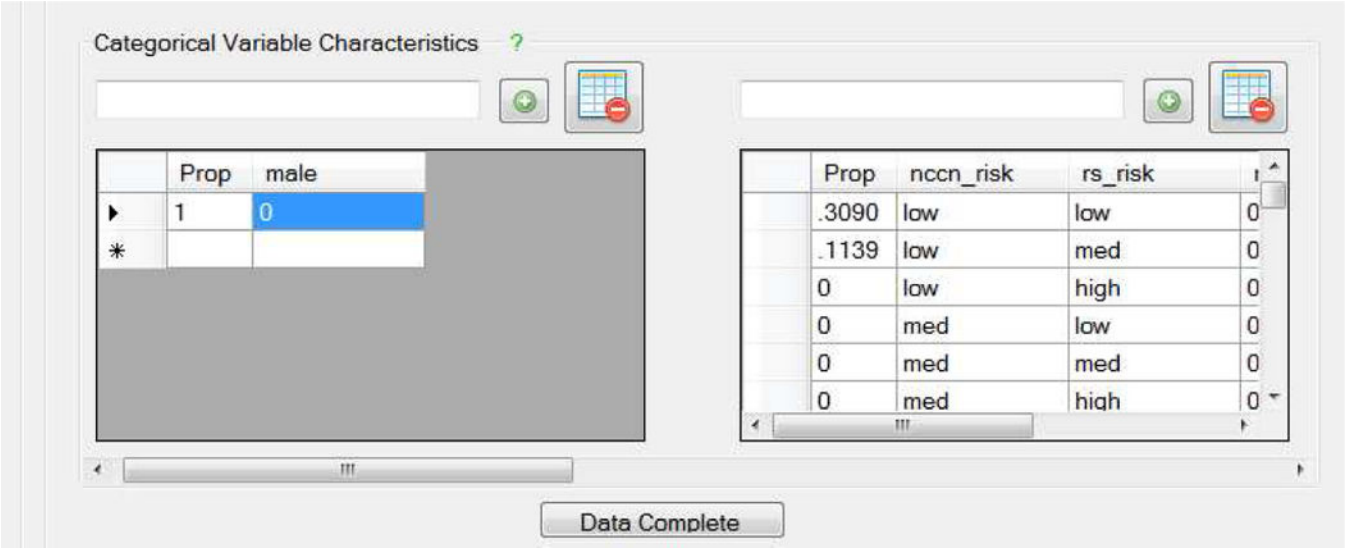
1. No chemo in either scenario (rx_notest=0 and rx_test=0)
2. Chemo without testing but no chemo with testing (rx_notest=1 and rx_test=0)
3. No chemo without testing but chemo with testing (rx_notest=0 and rx_test=1)
4. Chemo in both scenarios (rx_notest=1 and rx_test=1)

Prop	nccn_risk	rs_risk	rx_notest	rx_test
0.3090	low	low	0	0
0.1139	low	med	0	0
0	low	high	0	0
0	med	low	0	0

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Prop	nccn_risk	rs_risk	rx_notest	rx_test
0	med	med	0	0
0	med	high	0	0
0	high	low	0	0
0	high	med	0	0
0	high	high	0	0
0	low	low	1	0
0	low	med	1	0
0	low	high	1	0
0.0640	med	LOW	1	0
0.0137	med	med	1	0
0	med	high	1	0
0.0729	high	low	1	0
0.0233	high	med	1	0
0	high	high	1	0
0.0145	low	low	0	1
0.0119	low	med	0	1
0.0779	low	high	0	1
0	med	low	0	1
0	med	med	0	1
0	med	high	0	1
0	high	low	0	1
0	high	med	0	1
0	high	high	0	1
0	low	low	1	1
0	low	Med	1	1
0	low	high	1	1
0.0213	med	low	1	1
0.0223	med	med	1	1
0.0645	med	high	1	1
0.0243	high	low	1	1
0.0380	high	med	1	1
0.1285	high	high	1	1

7. When you are done filling out the table, press the “Data Complete” button under the table.



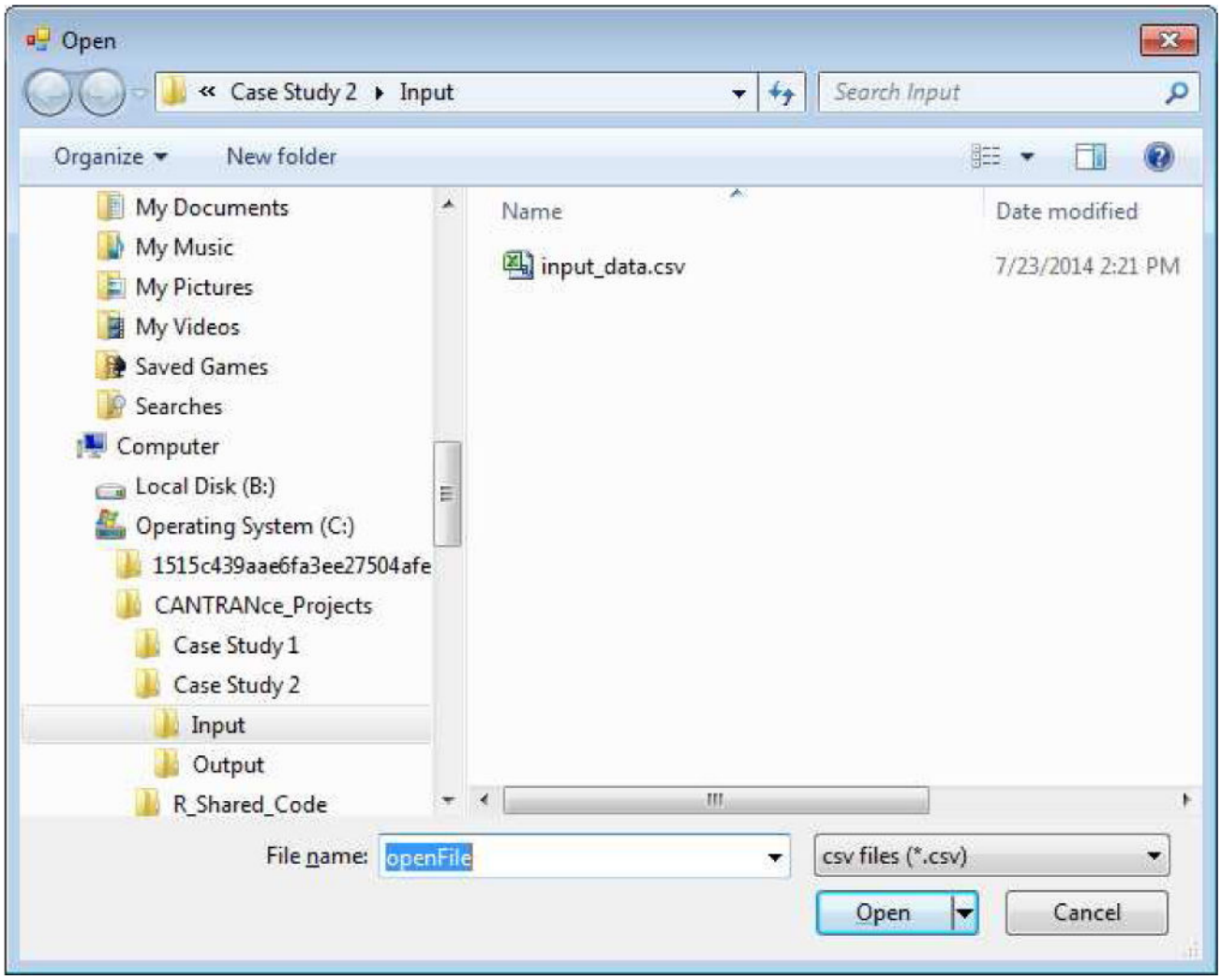
8. To continue running the Case 1 model, proceed to the section “Running the Case 1 or Case 2 model – Part II” on page 12.

Running the Case 2 model – Part I

1. When you choose “Individual-level Data” as your option from the Data menu, it will open a browser where you can choose a .csv file. For the Case 2 model, use the “input_data.csv” file that is located in the Input folder of the new project folder you created.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

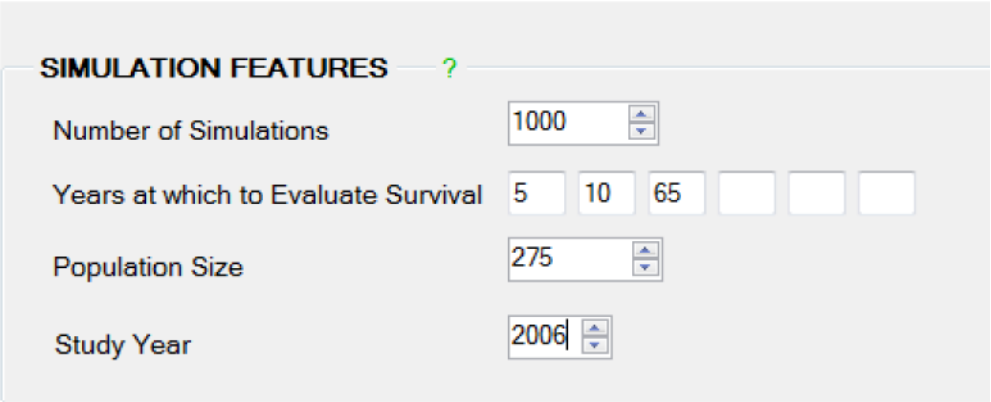
Author Manuscript Author Manuscript Author Manuscript Author Manuscript



This file is a public-access approximation to the the Ademuyiwa CE study data used in Case 2 (reference 14 in the manuscript). More details are given on the next page.

2. Enter the following parameters in the “Simulation Features” section.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Number of Simulations: 1000 simulations is a standard number of replicates to characterize uncertainty. Since we want to estimate lifetime results, **Years at which to Evaluate Survival** needs a maximum follow-up time that represents lifetime follow-up. Given that the average age was 55 in the Ademuyiwa CE study, we enter 65 years as the maximum follow-up because all women will die by age 120 according to our cohort life table. **Population Size** is 275, the number of subjects in the CE study, and **Study Year** is 2006, the year the CE data were collected.

3. The **Comparative Effectiveness Data** section is grayed out because all the data are contained in the input_data.csv file. The first few rows of the file are shown below.

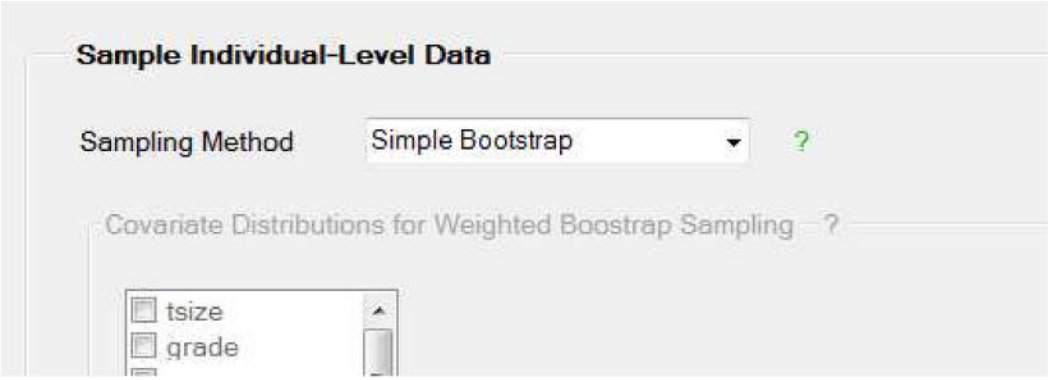
	A	B	C	D	E	F	G	H
1	tsize	grade	age	test_result	rx_notest	rx_test	male	
2	>2cm	1	40	low	1	0	0	
3	1.1-2cm	1	60	low	1	0	0	
4	<1cm	1	45	low	1	0	0	
5	>2cm	1	60	low	1	0	0	
6	1.1-2cm	1	65	low	1	0	0	
7	1.1-2cm	1	50	low	1	0	0	
8	<1cm	1	60	low	1	0	0	
9	1.1-2cm	1	65	low	1	1	0	

Each record represents a *simulated version of* one of the patients in the Ademuyiwa CE study data used in Case 2. All patients were female, so male=0 for all records. The individual values of the remaining covariates do not exactly match the true CE study data, but the population-level statistics for **age**, **test_result** (21-gene recurrence score risk group), **rx_notest** (treatment recommendation without testing, 1=chemo+hormone 0=hormone), and **rx_test** (treatment recommendation with testing) match those published in the Ademuyiwa et al paper, Tables 4 and 5 (see reference 14). The **tsize** and **grade** covariates are fake covariates you can play around with.

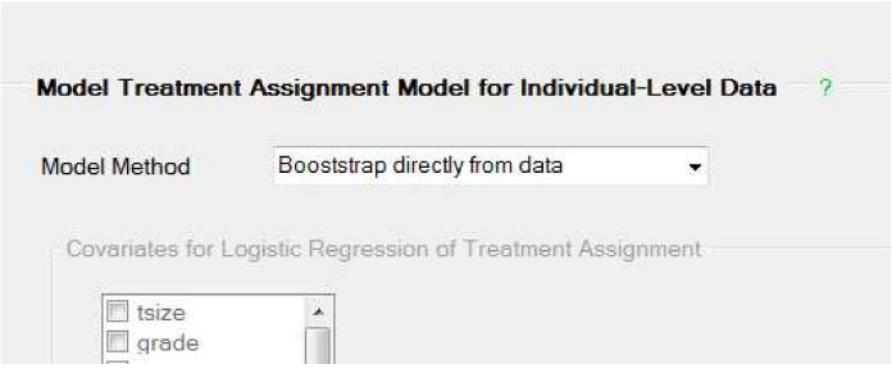
4. Leave **Sample Individual-Level Data** as “Simple Bootstrap.” This means that when the data is resampled for each simulation to represent sampling uncertainty, no weights will be used. Alternatively, a weighted bootstrap could be used to model a hypothetical population

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

with different covariate distributions (see the Methods section of the manuscript). For example, in the Case 2 data the 21-gene recurrence risk distribution is 51% low-risk, 40% medium-risk and 9% high-risk (see Table 2 in the manuscript). If you wanted to model outcomes for a higher-risk population, you could choose “Weighted Bootstrap” below, select “test_result” as the variable, and enter different proportions for the risk categories, e.g. 40% low-risk, 40% medium-risk, and 20% high-risk. The model would then resample the data, oversampling higher-risk cases and undersampling low-risk cases to match the distribution you specified. The remaining projection would pertain to the altered population.



5. Leave **Model Treatment Assignment for Individual-Level Data** as “Bootstrap directly from data.” This means that the model will take the values for treatment recommended in each scenario directly from the CE study data, rather than prompting the user for a logistic regression model from which to simulate treatment recommendations (see the Methods section of the manuscript for reasons why you might choose to do a logistic regression rather than bootstrap directly from data).

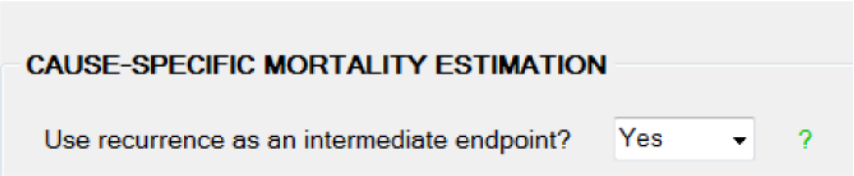


6. Continue to “Running the Case 1 or Case 2 model – Part II” on the next page.

Running the Case 1 or Case 2 model – Part II

1. Choose “Yes” to Use recurrence as an intermediate endpoint?

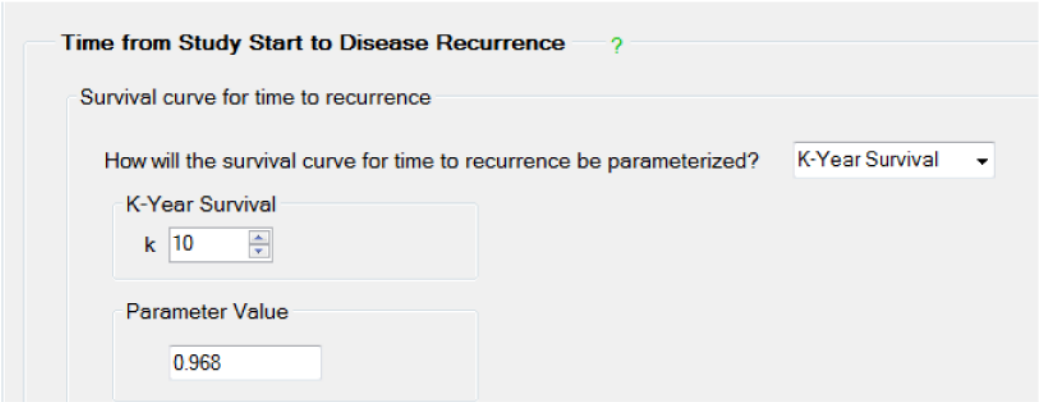
Author Manuscript Author Manuscript Author Manuscript Author Manuscript



In both models, we use recurrence as a second intermediate endpoint. This reflects the available long-term data on the particular diagnostic test intervention under study, the 21-gene recurrence score (RS). For this test, time from treatment to recurrence has been reported for risk subgroups defined by the test, but time from treatment to mortality has not. If time to mortality data were available, we would not have to use recurrence as an intermediate endpoint.

2. We will now enter the time to recurrence parameters from Table 1 of the manuscript.

Enter the parameters for **Survival curve for time to recurrence** as shown below. We choose the 10-year distant recurrence statistic for RS low risk women not receiving chemotherapy to initially define the time to recurrence survival curve. The curve is specified here defines the baseline curve to which subsequent hazard ratios will refer.

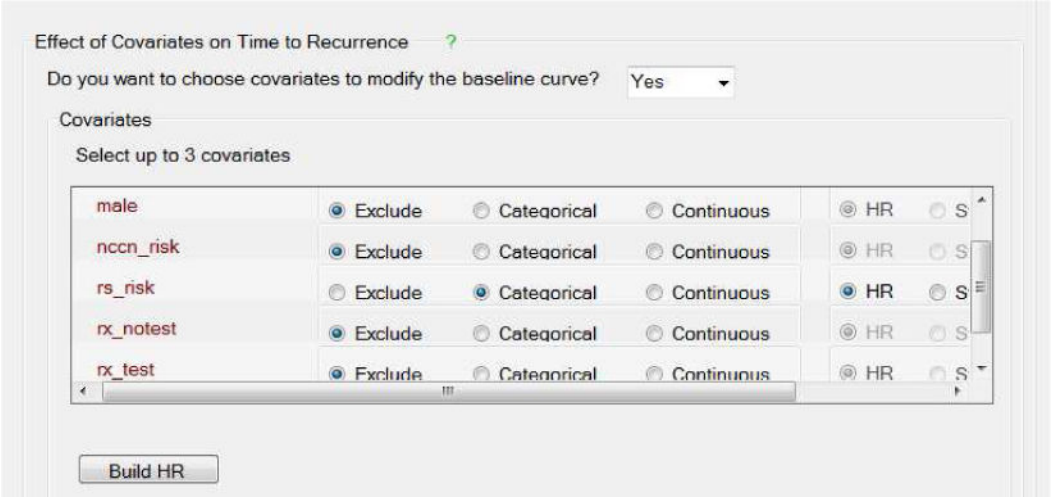


Note that this specification of a 10-year survival of 0.968 is particular to this specific application: it represents time to breast cancer recurrence for RS low risk women. In other words, 10 years after treatment, we expect 96.8% of RS low risk women who did not receive chemotherapy to be recurrence-free. The appropriate survival statistic will change if the model is applied to a different diagnostic test and/or different cancer.

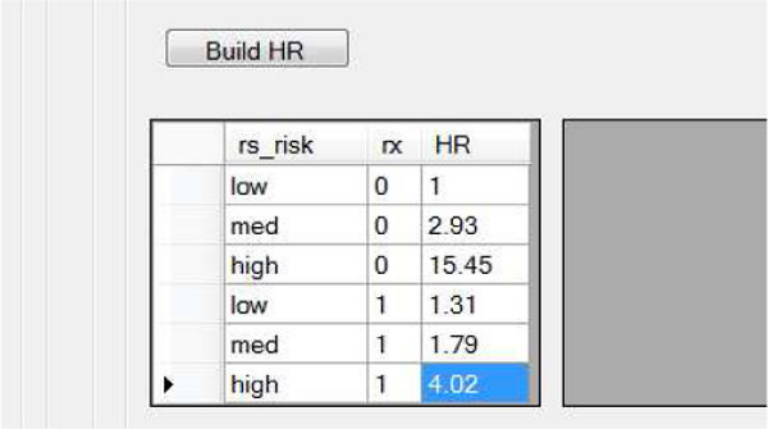
3. Now we will specify different times to recurrence for different RS risk and treatment groups. Choose “Yes” to **Do you want to choose covariates to modify the baseline curve?** Then, select “Categorical” and “HR” for the covariate “rs_risk” (Case 1) or “test_result” (Case 2), and click “Build HR.” We do this because Table 1 contains data on how time to recurrence varies by RS risk group and by treatment. While we primarily need to know how time to recurrence varies for tailored versus standard treatment, having these data stratified

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

by RS risk group can additionally help us be more precise in our projection. We will specify this variation in the form of hazard ratios (HRs).



4. Populate the table that is created as shown below. These HRs are derived from Table 1 in the paper.* The baseline group of RS low risk, no chemotherapy (rx=0) is indicating using HR=1. This is the group that we defined above as having a 10-year time to recurrence parameter of 0.968. All other hazard ratios are relative to this group. *Note: the covariate "rs_risk" shown below is called "test_result" in the Case 2 data, and the order of rows may be different. Please make sure you do not simply enter the HRs in order, but check that you are entering the correct HR for the risk and treatment group indicated in each row.*



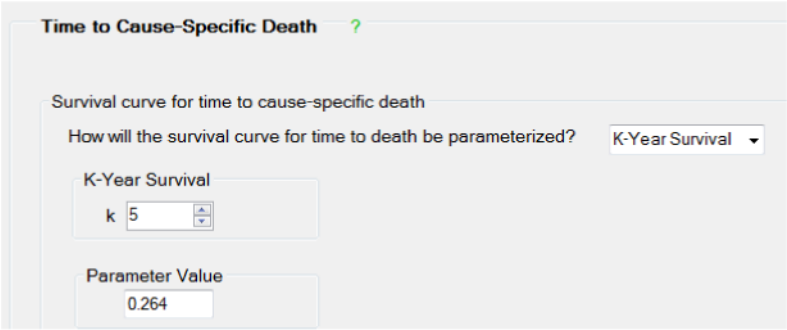
rs_risk	rx	HR
low	0	1
med	0	2.93
high	0	15.45
low	1	1.31
med	1	1.79
high	1	4.02

*Hazard ratios are the ratios of rates. We enter HRs here relative to the low-risk, no chemotherapy group, so the HRs are the ratios of each group's recurrence rate to that group. The Table 1 parameters are on the survival scale, so to compute the HRs we must transform back to the rate scale using rate = log(survival)/k, where k=the year at which survival was evaluated. For example, to compute the HR for med-risk, no chemo to low-risk, no chemo, we have the ratio of rates [log(0.909)/10]/[log(0.968)/10] = 2.93. To get the HR for med-risk

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

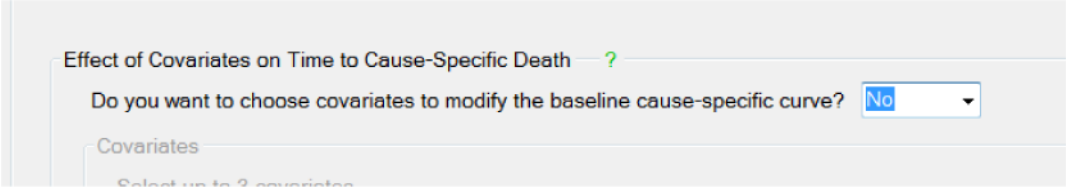
with chemo relative to low-risk, no chemo, we multiple the med-risk, chemo versus med-risk, no chemo with the HR we just computed, 0.61*2.93 = 1.79.

5. Enter the parameters for **Survival curve for time to cause-specific death** as shown below. Because we chose to go use disease recurrence as a second intermediate endpoint, this survival curve refers to the time from recurrence to mortality. If we had not chosen recurrence as a second intermediate endpoint, this time would reflect time from treatment to mortality.



As indicated in Table 1 of the manuscript, this 5-year survival of 0.264 is the relative breast cancer survival for women after distant recurrence of breast cancer. In other words, 5 years after distant recurrence, we expect 26.4% of women to have survived breast cancer death *net of death from other causes* (please see Methods and Discussion of manuscript regarding net survival).

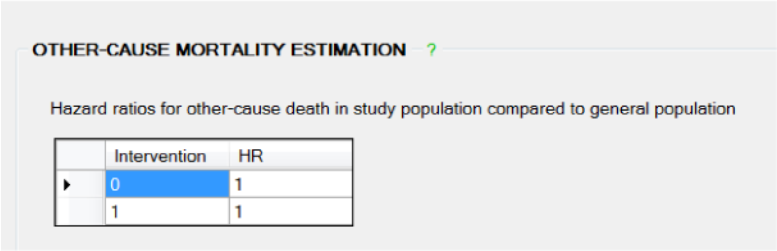
6. Choose “No” to **Do you want to choose covariates to modify the baseline cause-specific curve?** We do this because we assume that once patients recur, their remaining cause-specific survival does not vary significantly by any covariates. If we had evidence and data that available covariates significantly modified time from recurrence to mortality, we could choose “Yes” and enter hazard ratios for those covariates here.



7. Do not modify the **Other-Cause Mortality Estimation** section. We choose to make no assumptions about how other-cause (non-breast-cancer) mortality in our study population may differ from that of the general US population, which means that the hazard ratios (HRs) stay as 1. However, it is plausible that a study population may have a different overall survival profile from the US population. Trial populations, for example, may be healthier due to the selection criteria. Observational screening cohorts may be healthier due to the “healthy screenee effect” in which individuals who select to screen are a healthy subset of the general population. While it may be difficult to precisely quantify these types of effects,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

a HR lower than 1 could be used to explore the potential impact of the intervention on a healthier-than-US population.

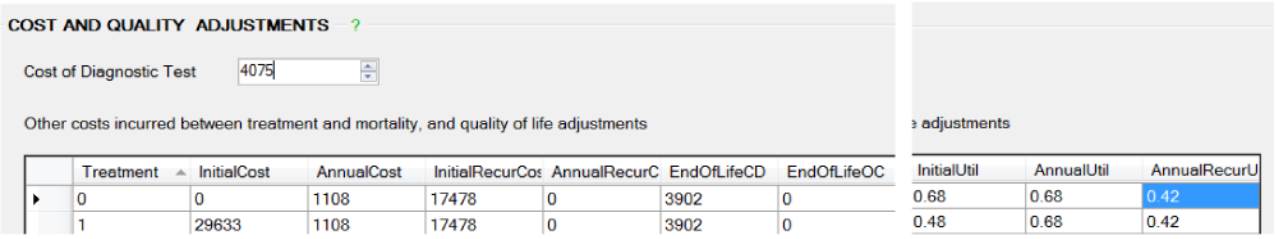


A NOTE ABOUT MODELING TIMES TO EVENT

As noted in the manuscript, CANTRANce models times to event using exponential survival curves. These curves assume a constant rate of event (e.g., recurrence or mortality), which is why you only had to enter one statistic to define the time to recurrence or time to mortality curve. While the constant rate assumption is convenient because it only requires one parameter, it may not always well-represent the true survival pattern. That being said, in many applications it will approximate true survival closely enough to produce informative results.

For a more extensive discussion of ways to represent survival, please see: *Connock M, Hyde C, Moore D. Cautions regarding the fitting and interpretation of survival curves: examples from NICE single technology appraisals of drugs for cancer. Pharmacoeconomics. 2011 Oct;29(10):827–837. PMID: 21770482*

8. Enter the following parameters in the **Cost and Quality Adjustments** section. Treatment=0 refers to patients who do not get chemotherapy (standard treatment), and Treatment=1 to those who do (tailored treatment). For Treatment=1, **InitialCost** is the sum of chemotherapy costs in the 1st year plus absence of work attributable to chemotherapy. All other costs follow directly from Table 1 of the manuscript. See the Help (hover over the green ?) for a detailed description of each column of the input table.



In the right panel, click “Check Parameters.” If no warning messages appear, you have entered all inputs without errors. Correct any errors until clicking “Check Parameters” results in no messages. Then, click “Run Analysis.”

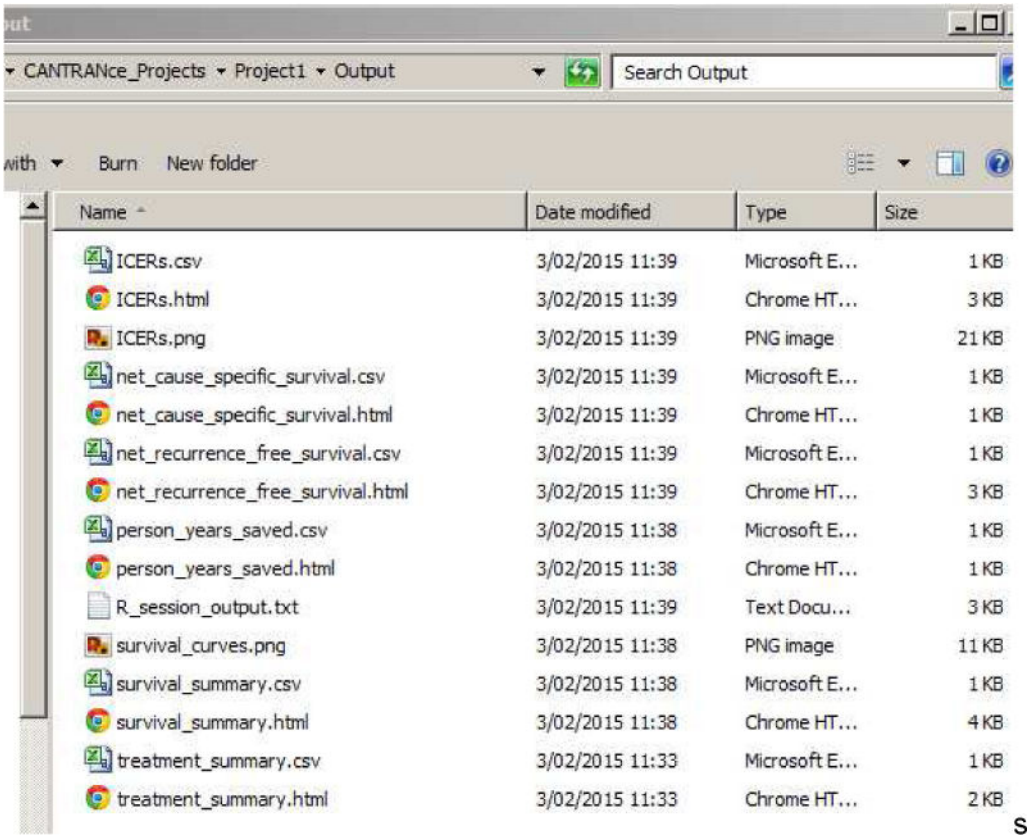
Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The model should take under 10 minutes to run for Case 1 and under 5 minutes for Case 2. If you receive any error messages after clicking “Run Analysis,” please follow the additional steps in the Troubleshooting section on page 25 to run the model.

Interpreting Results

After 10 minutes, you should see results in the Output folder, as in the screenshot below. If you do not, please follow the additional steps outlined in the Troubleshooting section on page 25 to run the model.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



S

This section will walk the user through interpreting the results from both Cases. Identically-named .csv and .html files contain the same data.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Outputs that validate Inputs

File: treatment_summary.csv/html

Mean and quantiles of treatment probability cross-classifications between scenarios (Group 0 = without diagnostic test; Group 1 = with diagnostic test)

Case 1

Group 0 Treatment	Group 1 Treatment	Mean	2.5%	50%	97.5%
0	0	0.423	0.392	0.423	0.455
0	1	0.104	0.086	0.104	0.123
1	0	0.173	0.151	0.173	0.199
1	1	0.299	0.271	0.299	0.329

Table 2 Treatment probabilities cross-classified by RS-guided and non-RS-guided strategies

	RS-guided strategy		
Non-RS-guided strategy	**Hormonal therapy**	**Chemo-therapy**	**Total**
Hormonal therapy, %	42.2	10.5	52.7
Chemotherapy, %	17.4	29.9	47.3
Total, %	59.6	40.4	100.0

RS, recurrence score.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Case 2

Group 0 Treatment	Group 1 Treatment	Mean	2.5%	50%	97.5%
0	0	0.430	0.374	0.429	0.487
0	1	0.123	0.084	0.124	0.164
1	0	0.266	0.215	0.265	0.316
1	1	0.181	0.138	0.182	0.229

Table 4 ODX-blinded recommendation and actual treatment received based on ODX score

	CTX received				All	
	No		Yes			
	N	%	*N*	%	*N*	%
ODX-blinded CTX recommendation						
No	117	42.3	34	**12.3**	151	54.7
Yes	71	**25.7**	54	19.7	125	45.3
All	188	68.0	88	32.0	276	100

The treatment_summary table shows the proportion of the virtual population falling into each of the four possible groups of treatment outcomes:

1) No chemo in either scenario (Group 0 Treatment=0 and Group 1 Treatment=0)
2) No chemo without testing but chemo with testing (Group 0 Treatment=0 and Group 1 Treatment=1)
3) Chemo without testing but no chemo with testing (Group 0 Treatment=1 and Group 1 Treatment=0)
4) Chemo in both scenarios (Group 0 Treatment =1 and Group 1 Treatment=1)

The 95% uncertainty intervals (2.5th percentile – 97.5th percentile) arise from the distribution across the 1000 simulations. An important feature of the model to understand is that each run of the model uses one parameter set, the parameters you entered, but that one parameter set is used for multiple simulations. As described in the Discussion, each simulation reflects the same population-level parameters, but the individual values of covariates are different for the same virtual person across simulations. The results are presented as a summary across simulations: the mean across simulations is the point estimate, and the 2.5% and 97.5% percentiles are used to construct a 95% uncertainty interval. We also sometimes present the median, the 50% percentile.

In Case 1, you can see that these proportions closely approximate those given in Reed Table 2 (RS-guided = Group 1 and Non-RS-guided = Group 0). For example, we simulate 0.423 or 42.3% receiving treatment=0, hormonal therapy, in both scenarios, and below, Reed shows 42.2%. We simulate 0.299 or 29.9% receiving chemo in both scenarios (treatment=1), exactly as Reed does. Similarly, the virtual population in Case 2 closely approximates Ademuyiwa Table 4 (CTX received = Group 1 and ODX-blinded recommendation = Group 0). For example, both indicate 12.3% of patients were switched onto chemotherapy with testing.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

File: net_recurrence_free_survival.csv/html

10-year recurrence-free survival in the simulated population in the absence of other-cause death

Case 1

Covariate	Value	Treatment	K-Time Survival	2.5%	50%	97.5%
All	All	All	0.927	0.920	0.927	0.934
All	All	0	0.937	0.928	0.936	0.945
All	All	1	0.914	0.903	0.914	0.925
rs_risk	low	0	0.968	0.960	0.968	0.975
rs_risk	med	0	0.909	0.889	0.909	0.929
rs_risk	high	0	0.605	0.515	0.605	0.681
rs_risk	low	1	0.958	0.942	0.958	0.973
rs_risk	med	1	0.943	0.921	0.943	0.965
rs_risk	high	1	0.878	0.859	0.878	0.896

Case 2

Covariate	Value	Treatment	K-Time Survival	2.5%	50%	97.5%
All	All	All	0.940	0.926	0.940	0.952
All	All	0	0.943	0.928	0.943	0.958
All	All	1	0.935	0.913	0.936	0.956
test_result	high	0	0.581	0.139	0.600	0.920
test_result	low	0	0.968	0.953	0.968	0.982
test_result	med	0	0.909	0.873	0.910	0.940
test_result	high	1	0.875	0.796	0.878	0.939
test_result	low	1	0.958	0.917	0.960	0.991
test_result	med	1	0.943	0.914	0.944	0.967

The net_recurrence_free_survival table calculates the rate of recurrence in the virtual population. The table uses the same statistic type inputted by the user, and reports recurrence in the whole population ("All", "All", "All" row), by treatment status ("All", "All", 0/1 rows), and by the groups by which the user inputted hazard ratios (HRs) or different survival statistics (remaining rows). In both case studies, we put in a baseline 10-year survival statistic and hazard ratios by the six RS-treatment groups:

1) low, no chemo (treatment=0)
2) med, no chemo
3) high, no chemo
4) low, chemo (treatment=1)
5) med, chemo
6) high, chemo

The table thus reports 10-year survival statistics for recurrence for each of these groups. You can see that the low-0 row (RS low-risk, no chemo) has a statistic of 0.968, the baseline survival we put in. The other statistics must be converted back to HRs to be compared to our inputs. This must be done on the log scale, e.g. the HR for med-0 (RS med-risk, no chemo) in Case 1 is log(0.909)/log(0.968) = 2.93, exactly what we put in as the HR.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

File: net_cause_specific_survival.csv/html

5-year cause-specific survival in the simulated population in the absence of other-cause death

Case 1

Covariate	Value	Treatment	K-Time Survival	2.5%	50%	97.5%
All	All	All	0.263	0.226	0.263	0.302
All	All	0	0.263	0.206	0.263	0.315
All	All	1	0.263	0.208	0.263	0.317

Case 2

Covariate	Value	Treatment	K-Time Survival	2.5%	50%	97.5%
All	All	All	0.263	0.178	0.265	0.349
All	All	0	0.262	0.158	0.263	0.372
All	All	1	0.264	0.136	0.265	0.402

The net_cause_specific_survival table calculates the rate of cause-specific death in the virtual population. When recurrence is used as an intermediate endpoint, as in both case studies, this refers to the time from recurrence to cause-specific death. The format of this table is similar to that of the previous table. The type of statistic reported again corresponds to that input by the user. Since we inputted 5-year survivals, that is the output we see, and the results approximate the value of 0.264 we put in.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Outputs that present results

File: survival_summary.csv

Mean, median, and k-year all-cause, crude, and net survival by testing scenario (Group 0 = without diagnostic test; Group 1 = with diagnostic test)

	Measure	Group	Mean Survival	Lower	Upper	Median Survival	Lower	Upper	5-Year Survival	Lower	Upper	10-Year Survival	Lower	Upper	65-Year Survival	Lower	Upper
Case 1	All-Cause Survival	0	25.2	24.5	25.9	26.0	24.9	27.1	0.948	0.934	0.961	0.865	0.843	0.886	0.000	0.000	0.000
	All-Cause Survival	1	25.8	25.0	26.5	26.7	25.6	27.8	0.954	0.940	0.966	0.879	0.858	0.900	0.000	0.000	0.000
	Net Survival	0	192.6	178.3	208.0	107.4	95.8	119.2	0.978	0.969	0.987	0.940	0.925	0.954	0.634	0.604	0.664
	Net Survival	1	203.9	188.7	220.4	115.8	104.5	127.3	0.984	0.976	0.992	0.955	0.942	0.967	0.668	0.639	0.698
	Crude Survival	0							0.978	0.970	0.987	0.942	0.928	0.956	0.818	0.793	0.841
	Crude Survival	1							0.984	0.976	0.992	0.957	0.944	0.969	0.848	0.825	0.870
	Measure	**Group**	**Mean Survival**	**Lower**	**Upper**	**Median Survival**	**Lower**	**Upper**	**5-Year Survival**	**Lower**	**Upper**	**10-Year Survival**	**Lower**	**Upper**	**65-Year Survival**	**Lower**	**Upper**
Case 2	All-Cause Survival	0	23.7	22.0	25.3	22.3	20.1	24.5	0.922	0.889	0.953	0.817	0.770	0.860	0.003	0.000	0.008
	All-Cause Survival	1	23.7	22.0	25.4	22.3	20.1	24.6	0.922	0.889	0.953	0.817	0.768	0.860	0.003	0.000	0.008
	Net Survival	0	202.2	176.9	230.8	123.8	103.2	147.1	0.984	0.969	1.000	0.957	0.931	0.979	0.687	0.634	0.742
	Net Survival	1	213.0	184.8	242.8	128.0	105.6	153.3	0.984	0.968	0.995	0.957	0.933	0.980	0.694	0.639	0.746
	Crude Survival	0							0.985	0.970	0.997	0.961	0.937	0.981	0.874	0.833	0.909
	Crude Survival	1							0.985	0.970	0.997	0.961	0.937	0.982	0.877	0.836	0.913

This table presents three types of survival statistics ("Measure" column): all-cause, net and crude survival. All-cause survival is overall survival, i.e. survival from all causes of death. Net and crude survival are different ways of measuring cause-specific death. Net survival censors other causes of death, effectively creating a hypothetical world in which everyone dies of the cancer eventually. Crude survival removes those who have died from other causes of death from the at-risk population, so survival plateaus once everyone has died of other causes. Mean and median crude survivals do not exist for the above projections because survival plateaus before even 50% of the population dies. Crude survival can be thought of as the inverse of cumulative incidence.

The different summaries of each statistic are the mean (average time to death across all individuals), median (time at which 50% of the population is still alive), and proportion of people surviving at the time points specified by the user in the Simulation Features section of the interface. The "Lower" and "Upper" columns represent the 95% uncertainty interval across simulations.

Some highlights of these results: for both the 5- and 10-year Case 1 survivals, the model projects that testing saves 1 life per every 100 patients (1% better survival in the with-testing group). The 65-year survival is a lifetime estimate that suggests that over the lifetime of the whole Case 1 population, about 3 breast cancer deaths per 100 cases are averted by testing. In the Case 2 model, survival is virtually identical for the with- and without-testing groups by all measures, except for long-term net survival.

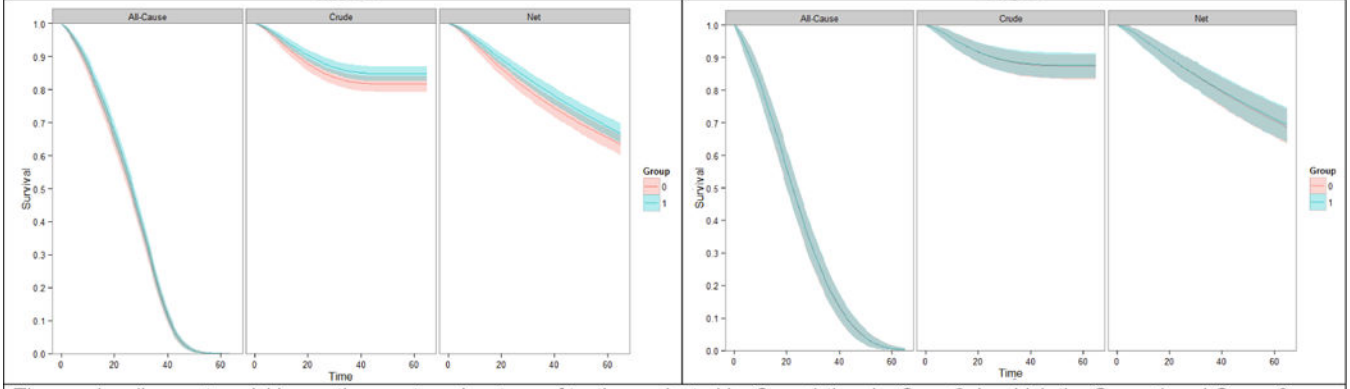
These numeric summaries are represented graphically in the survival_summary.png file, shown below.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

File: survival_summary.png

All-cause, crude and net survival from time of treatment (Time=0) to the maximum follow-up time specified. The line plotted is the median across simulations and the uncertainty bounds represent the 2.5% and 97.5% quantiles.

Case 1 | **Case 2**



The graphs allow us to quickly see the greater advantage of testing projected by Case 1 than by Case 2, in which the Group 1 and Group 0 curves overlap almost entirely. There is also greater uncertainty across simulations in Case 2, due to the smaller population size (N=275) compared to Case 1 (N=1000).

File: person_years_saved.csv/html

Total life-years saved by testing (non-discounted)

Case 1		Case 2	
Mean	585	Mean	15
Lower Bound	7	Lower Bound	-262
Upper Bound	1179	Upper Bound	268

This table gives the non-discounted, total number of life-years saved by testing over the maximum follow-up time, which we set as 65 years. This is a population-level statistic that will reflect the size of the population, i.e. larger populations will have more potential for more absolute life-years saved.

File: ICERS.csv/html

Per-person life-years, quality-adjusted life-years (QALYs), costs, and cost-effectiveness ratios (discounted at 3%)

		Group 0 Life-Years	Group 1 Life-Years	Group 0 QALYs	Group 1 QALYs	Group 0 Costs	Group 1 Costs	Incremental Life-Years	Incremental QALYs	Incremental Cost	ICER for Life-Years	ICER for QALYs
Case 1	Mean	16.96	17.28	11.32	11.58	34106	36131	0.32	0.25	2025	6361	8008
	Lower	16.58	16.86	11.06	11.29	33103	35139	0.01	0.03	1036	73140	29843
	Median	16.97	17.29	11.33	11.58	34105	36127	0.32	0.25	2016	6315	8018
	Upper	17.32	17.67	11.58	11.85	35082	37153	0.63	0.48	2999	4763	6311
Case 2		Group 0 Life-Years	Group 1 Life-Years	Group 0 QALYs	Group 1 QALYs	Group 0 Costs	Group 1 Costs	Incremental Life-Years	Incremental QALYs	Incremental Cost	ICER for Life-Years	ICER for QALYs
	Mean	15.90	15.92	10.64	10.69	31422	31246	0.02	0.04	-176		
	Lower	15.05	15.05	10.05	10.10	29586	29346	-0.47	-0.31	-2237		
	Median	15.90	15.92	10.65	10.69	31411	31230	0.02	0.05	-188		
	Upper	16.74	16.77	11.22	11.27	33426	33054	0.51	0.39	1839	3633	4721

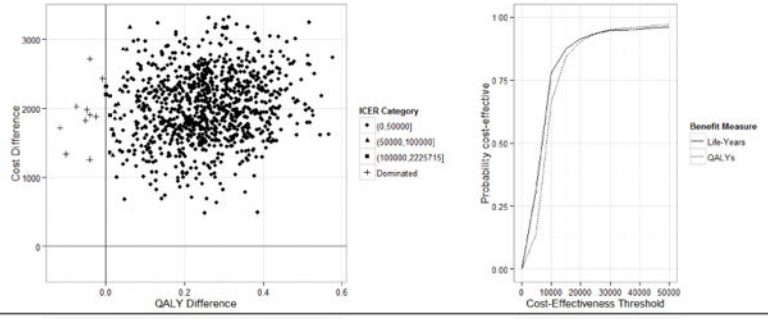
These results approximate the Case 1 and Case 2 results reported in the manuscript Table 3. For Case 1, the small differences arise from the truncated decimals used when hand-keying in the table of proportions. For Case 2, the differences arise from using a simulated version of the true CE data.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

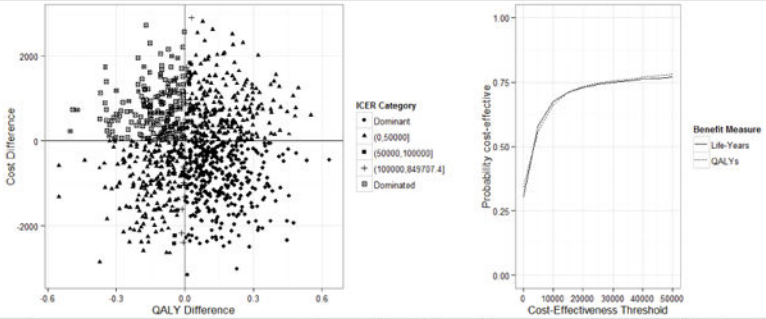
File: ICERS.png

Cost-effectiveness plane (left) and cost-effectiveness acceptability curve (right).

Case 1



Case 2



Each point on the cost-effectiveness plane represents the incremental cost (Y-axis) and incremental QALYs (X-axis) of with-testing versus without-testing *for one simulation*. Since we ran 1000 simulations in both cases, both graphs have 1000 points. As described in the Discussion of the manuscript, each simulation in the microsimulation model incorporates variation within the population. The wide range across simulations that all use the same parameter set comes from the individual-level variation in covariate values across simulations. The cost-effectiveness acceptability curves show that of the 1000 simulations, almost all Case 1 ICERs and about 75% of Case 2 ICERs fall below $50K/QALY.

Troubleshooting

Clicking “setup.exe” does not start CANTRANce.

Use the Program Manager to uninstall CANTRANce, and then re-install using “setup.exe.”

After clicking “Run Analysis,” the viewport to R stays blank but you see an “Analysis completed” message.

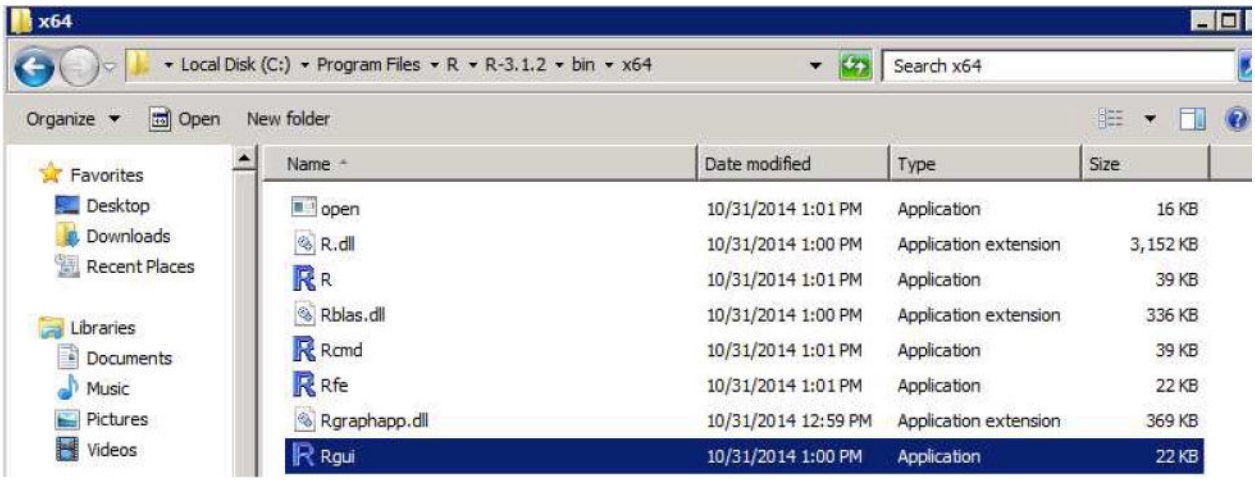
Correct Settings -> Path to R. The location for R on your machine is not correct.

After clicking “Run Analysis,” the viewport to R gives a “path not found” error, or the viewport stays blank and the model appears to run for more than 10 minutes without any files appearing in the C:\CANTRANce_Projects\Project1\Output folder.

Execute the following additional steps to run the model:

1. Open “Rgui.exe”, which is accessible from the location you installed R (page 1). It is most likely in the folder “C:\Program Files\R\R-3.1.2\bin\x64”.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



2. In the console that appears, type or copy-paste:

```
setwd("C:\\CANTRANce_Projects\\Project1\\")
```

and hit Enter. Please note the use of double backslashes. Project1 refers to the folder name you chose for your project in step 7 of “Installing and initializing the Diagnostics model interface, and accessing R code” (page 1).

```
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> setwd("C:\\CANTRANce_Projects\\Project1")
>
```

3. Type or copy-paste:

```
source("Input\\run_file.r")
```

and hit Enter. Again, note the double backslashes. It will look like this in the console:

```
> source("Input\\run_file.r")
```

4. The model will now run. You should see the results saved in your C:\CANTRANce_Projects\Project1\Output folder in 5–10 minutes.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Other problems

Close CANTRANce and re-start the application. If the issue remains unresolved, contact kurian@uw.edu.

References

1. Bates S. Progress towards personalized medicine. Drug Discov Today. 2010; 15:115–120. [PubMed: 19914397]
2. Chung C, Christianson M. Predictive and prognostic biomarkers with therapeutic targets in breast, colorectal, and non-small cell lung cancers: A systemic review of current development, evidence, and recommendation. J Oncol Pharm Pract. 2013; doi: 10.1177/1078155212474047
3. Workman P. The opportunities and challenges of personalized genome-based molecular therapies for cancer: targets, technologies, and molecular chaperones. Cancer Chemother Pharmacol. 2003; 52(Suppl 1):S45–56. [PubMed: 12819933]
4. Ow TJ, Sandulache VC, Skinner HD, Myers JN. Integration of cancer genomics with treatment selection: From the genome to predictive biomarkers. Cancer. 2013; doi: 10.1002/cncr.28304
5. Chung C, Christianson M. Predictive and prognostic biomarkers with therapeutic targets in breast, colorectal, and non-small cell lung cancers: A systemic review of current development, evidence, and recommendation. J Oncol Pharm Pract. 2014; 20:11–28. [PubMed: 23493335]
6. Rizzo S, et al. Prognostic vs predictive molecular biomarkers in colorectal cancer: is KRAS and BRAF wild type status required for anti-EGFR therapy? Cancer Treat Rev. 2010; 36(Suppl 3):S56–61. [PubMed: 21129611]
7. Linardou H, et al. Assessment of somatic k-RAS mutations as a mechanism associated with resistance to EGFR-targeted agents: a systematic review and meta-analysis of studies in advanced non-small-cell lung cancer and metastatic colorectal cancer. Lancet Oncol. 2008; 9:962–972. [PubMed: 18804418]
8. Wang SC, Zhang L, Hortobagyi GN, Hung MC. Targeting HER2: recent developments and future directions for breast cancer patients. Semin Oncol. 2001; 28:21–29. [PubMed: 11774202]
9. Bravatà V, Cammarata FP, Forte GI, Minafra L. 'Omics' of HER2-positive breast cancer. OMICS. 2013; 17:119–129. [PubMed: 23421906]
10. Dent S, Oyan B, Honig A, Mano M, Howell S. HER2-targeted therapy in breast cancer: a systematic review of neoadjuvant trials. Cancer Treat Rev. 2013; 39:622–631. [PubMed: 23434074]
11. Azim HA, et al. Utility of prognostic genomic tests in breast cancer practice: The IMPAKT 2012 Working Group Consensus Statement. Ann Oncol. 2013; 24:647–654. [PubMed: 23337633]
12. Lo SS, et al. Prospective multicenter study of the impact of the 21-gene recurrence score assay on medical oncologist and patient adjuvant breast cancer treatment selection. J Clin Oncol. 2010; 28:1671–1676. [PubMed: 20065191]
13. Joh JE, et al. The effect of Oncotype DX recurrence score on treatment recommendations for patients with estrogen receptor-positive early stage breast cancer and correlation with estimation of recurrence risk by breast cancer specialists. Oncologist. 2011; 16:1520–1526. [PubMed: 22016474]
14. Ademuyiwa FO, et al. The effects of oncotype DX recurrence scores on chemotherapy utilization in a multi-institutional breast cancer cohort. Breast Cancer Res Treat. 2011; 126:797–802. [PubMed: 21197567]
15. Müller BM, et al. The EndoPredict Gene-Expression Assay in Clinical Practice-Performance and Impact on Clinical Decisions. PLoS ONE. 2013; 8:e68252. [PubMed: 23826382]
16. Webster J, et al. KRAS testing and epidermal growth factor receptor inhibitor treatment for colorectal cancer in community settings. Cancer Epidemiol Biomarkers Prev. 2013; 22:91–101. [PubMed: 23155138]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

17. Torrisi R, et al. Potential impact of the 70-gene signature in the choice of adjuvant systemic treatment for ER positive, HER2 negative tumors: a single institution experience. Breast. 2013; 22:419–424. [PubMed: 23643803]
18. Glasgow RE, et al. Comparative effectiveness research in cancer: what has been funded and what knowledge gaps remain? J Natl Cancer Inst. 2013; 105:766–773. [PubMed: 23578853]
19. Lyman GH, Levine M. Comparative effectiveness research in oncology: an overview. J Clin Oncol. 2012; 30:4181–4184. [PubMed: 23071249]
20. Rutter CM, Zaslavsky AM, Feuer EJ. Dynamic microsimulation models for health outcomes: a review. Med Decis Making. 2011; 31:10–18. [PubMed: 20484091]
21. Siebert U, et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. Med Decis Making. 2012; 32:690–700. [PubMed: 22990084]
22. Briggs AH, et al. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6. Med Decis Making. 2012; 32:722–732. [PubMed: 22990087]
23. Cronin KA, Feuer EJ. Cumulative cause-specific mortality for cancer patients in the presence of other causes: a crude analogue of relative survival. Stat Med. 2000; 19:1729–1740. [PubMed: 10861774]
24. National Cancer Institute. SEER Cancer Statistics Review (CSR) 1975–2011. 2014. at <http://seer.cancer.gov/csr/1975_2011/>
25. National Cancer Institute. Surveillance, Epidemiology and End Results Program. at <www.seer.cancer.gov>
26. The Berkeley Mortality Database. at <http://demog.berkeley.edu/~bmd/>
27. Pinsky PF, et al. Evidence of a healthy volunteer effect in the prostate, lung, colorectal, and ovarian cancer screening trial. Am J Epidemiol. 2007; 165:874–881. [PubMed: 17244633]
28. Krahn M, Gafni A. Discounting in the economic evaluation of health care interventions. Med Care. 1993; 31:403–418. [PubMed: 8501989]
29. Gold, MR.; Siegel, JE.; Russell, LB.; Weinstein, MC. Cost-Effectiveness in Health and Medicine. Oxford University Press; 1996.
30. Reed SD, Dinan MA, Schulman KA, Lyman GH. Cost-effectiveness of the 21-gene recurrence score assay in the context of multifactorial decision making to guide chemotherapy for early-stage breast cancer. Genet Med. 2012; doi: 10.1038/gim.2012.119
31. Tang G, et al. Comparison of the prognostic and predictive utilities of the 21-gene Recurrence Score assay and Adjuvant! for women with node-negative, ER-positive breast cancer: results from NSABP B-14 and NSABP B-20. Breast Cancer Res Treat. 2011; 127:133–142. [PubMed: 21221771]
32. Fleming TR, Prentice RL, Pepe MS, Glidden D. Surrogate and auxiliary endpoints in clinical trials, with potential applications in cancer and AIDS research. Stat Med. 1994; 13:955–968. [PubMed: 8047747]
33. Gill S, Sargent D. End points for adjuvant therapy trials: has the time come to accept disease-free survival as a surrogate end point for overall survival? Oncologist. 2006; 11:624–629. [PubMed: 16794241]
34. Jatoi I, Anderson WF, Jeong JH, Redmond CK. Breast cancer adjuvant therapy: time to consider its time-dependent effects. J Clin Oncol. 2011; 29:2301–2304. [PubMed: 21555693]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

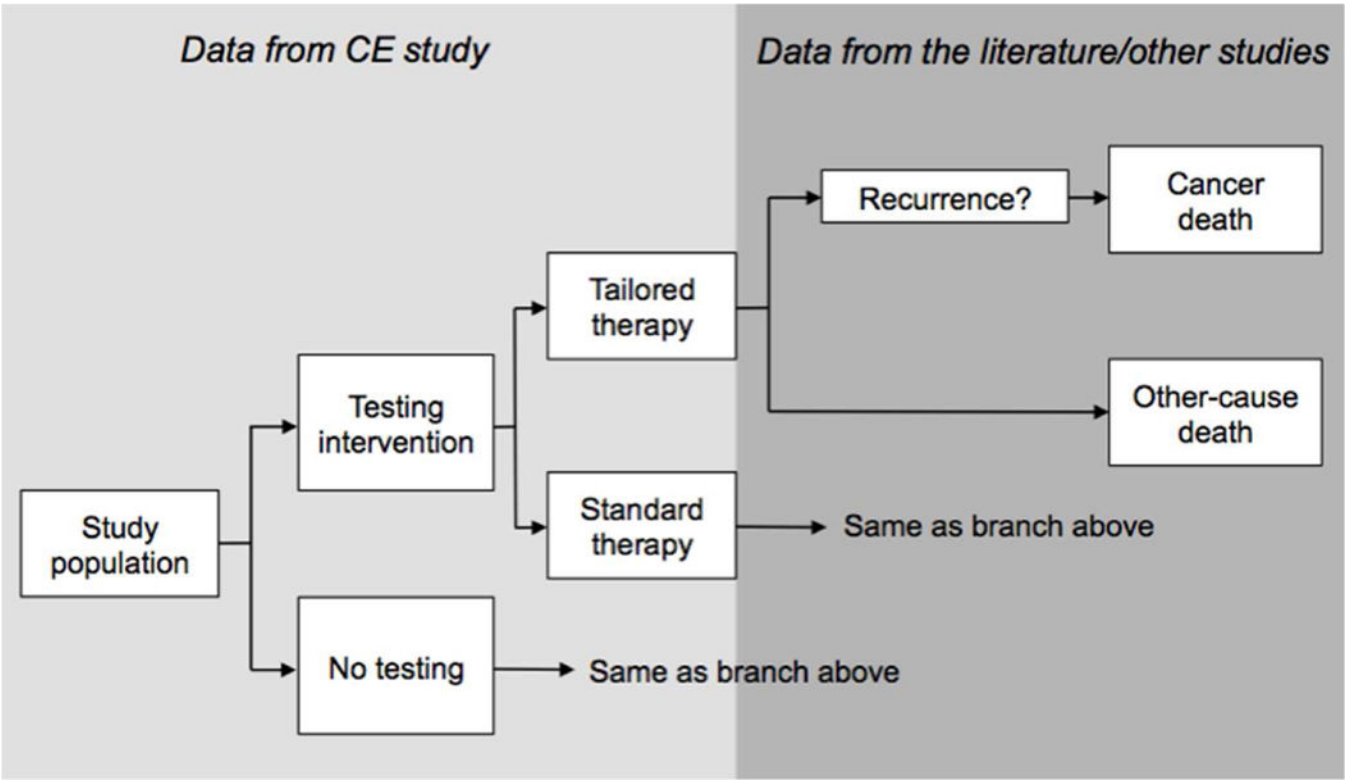


Figure 1.

Schematic of the diagnostics model. The virtual study population is subjected to two scenarios: all are tested for the diagnostic biomarker, or none are tested. The resulting distribution of tailored versus standard therapy for each scenario is based on a CE study. CANTRANce uses additional data from the literature/other studies to project from the time of treatment to mortality.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

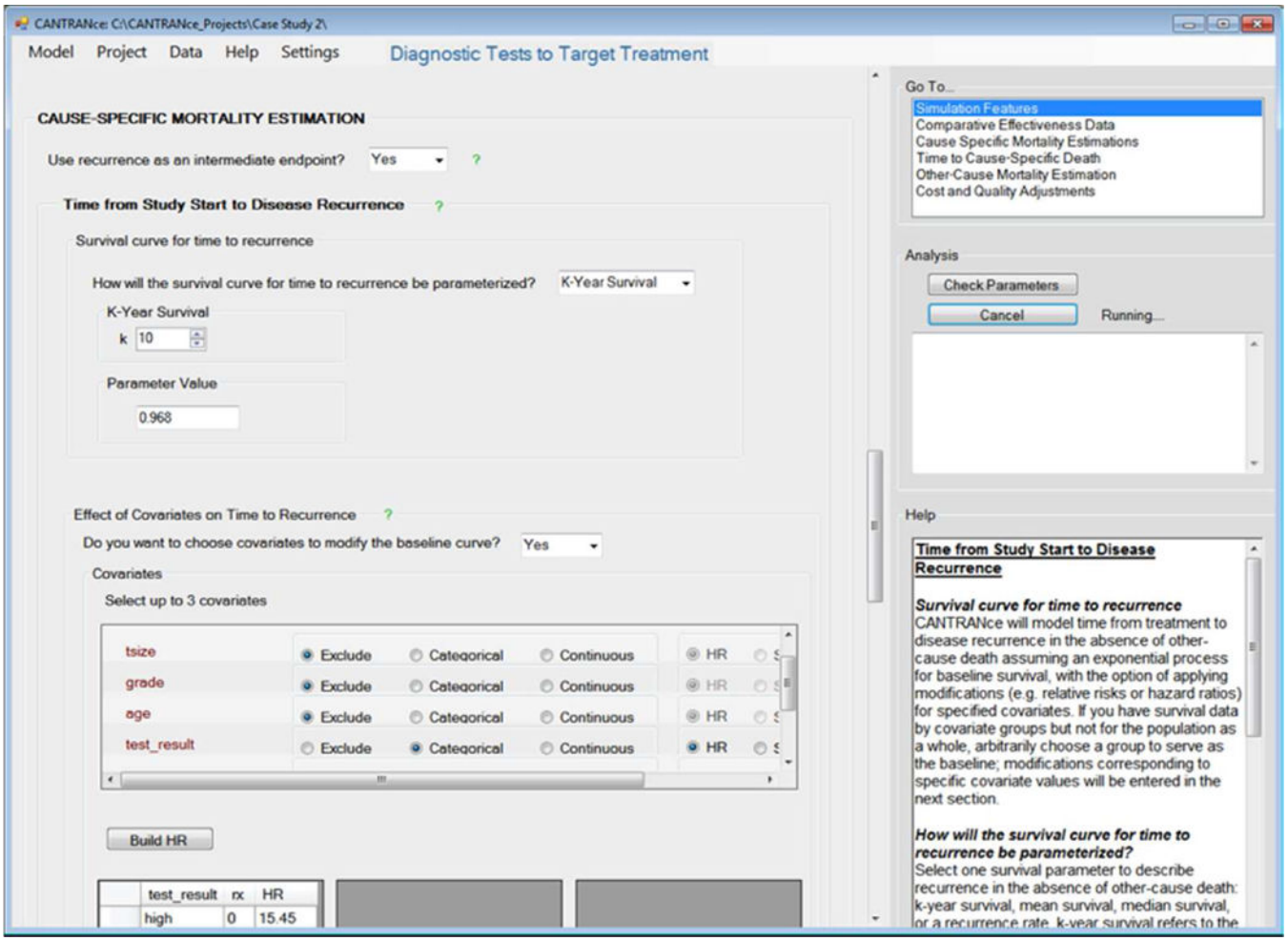


Figure 2.
Screenshot of the user interface for the diagnostics model.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1

Diagnostics Model Parameters for the Two Case Studies, Based on Reed et al[30]

Parameter	Mean (SE)
10-year distant recurrence-free survival with hormone therapy[a]	
RS low risk	0.968 (0.016)
RS intermediate risk	0.909 (0.043)
RS high risk	0.605 (0.073)
Hazard ratio of chemotherapy on distant recurrence[b]	
RS low risk	1.31 (0.57)
RS intermediate risk	0.61 (0.56)
RS high risk	0.26 (0.31)
5-year relative survival after distant recurrence	0.264
Health state utilities	
Chemotherapy in the first year	0.48 (0.06)
Hormonal therapy/remission	0.68 (0.06)
Distant recurrence	0.42 (0.06)
Medical costs, 2011 USD[c]	
Costs incurred in the first year	
21-gene recurrence score assay	4075
Chemotherapy in the first year	16,947 (1,655)
Absence from work attributable to chemotherapy	12,686
Monitoring and follow-up during remission[d]	1,108 (61)
Distant recurrence, one-time cost	17,478 (2,444)
Patient time during last year of life with metatstatic breast cancer	3,902

Distributional assumptions for PSA (probabilistic sensitivity analysis) also follow from Reed et al(30): beta distributions for probabilities and utilities, normal distributions for costs and log-relative risks, and Dirichlet for the multinomial parameters describing the CE study covariate summaries (not shown).

[a]This is a composite statistic used to approximate the net statistic of time to recurrence.

[b]Standard error is on the log scale.

[c]5-year annual costs of $105 for hormonal therapy, included by Reed, were not included.

[d]These costs were applied until recurrence occurred, whereas Reed et al limited them to 10 years.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2

Distribution of Patients by RS and Percent of Patients Receiving Chemotherapy (CTX) within RS Groups (Low, Intermediate, and High), for Each Case Study

	Case 1			Case 2		
	RS Distribution[a]	CTX without RS[b]	CTX with RS[b]	RS Distribution[a]	CTX without RS[b]	CTX with RS[b]
Low	53%	42%	13%	51%	37%	9%
Intermediate	19%	50%	36%	40%	47%	47%
High	29%	67%	100%	9%	88%	96%

Data for Case 1 come from Tang et al[31] and Lo et al (N=89).[12] Data for Case 2 come from Ademuyiwa et al (N=276).[14]

[a]Column % across RS groups.

[b]Row % within RS group.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 3

Lifetime Per-Patient Outcomes for the without RS Scenario, and Incremental Outcomes Comparing with RS to without RS, with 95% Uncertainty Intervals in Parentheses

	Reed Model	Case 1 – Sensitivity	Case 1	Case 2
Life-years *without RS*	14.82 (14.46–15.07)	16.02 (14.19–16.71)	16.98 (14.22–18.15)	17.12 (16.19–17.69)
Incremental life-years *with RS*	0.19 (0.09–0.32)	0.23 (−0.13–1.69)	0.33 (−0.23–2.33)	0.05 (−0.03–0.30)
QALYs *without RS*	9.93 (8.12–11.60)	10.74 (8.70–12.56)	11.34 (8.95–13.43)	11.48 (9.42–13.33)
Incremental QALYs *with RS*	0.16 (0.08–0.28)	0.18 (−0.04–1.19)	0.26 (−0.11–1.77)	0.06 (0.00–0.25)
Total costs *without RS*	24,656 (22,599–26,887)	31,943 (17,238–49,269)	34,172 (19,317–51,659)	33,153 (30,658–35,556)
Incremental total costs *with RS*	1,741 (−85–3,710)	1,939 (−15,757–21,363)	1,920 (−15,943–21,291)	26 (−412–482.68)

Reed Model results come from reference 29. Case 1 – Sensitivity replicates Reed including the life table and recurrence assumption, while Case 1 replicates Reed using the standard Diagnostics model life table and recurrence assumption. Case 2 is a novel projection of the mortality impact of testing implied by its impact on treatment in Ademuyiwa et al.[14] QALYs = quality-adjusted life-years.