User guide for the IPDfromKM Shiny app

Yanhong Zhou, Na Liu, J. Jack Lee

1 Introduction

In addition to the *IPDfromKM* R package, to facilitate the use of the modified-iKM algorithm improved upon the iKM approach [2] to reconstruct IPD, we develop the IPDfromKM Shiny app, which is freely available at `www.trialdesign.org`.

The IPDfromKM Shiny app is used to reconstruct individual patients data (IPD), i.e., time and event status from K-M curves. There are two main steps to obtain the IPD as below:

1. Extract raw data coordinates from published K-M curves.

2. Reconstruct the IPD from the data coordinates, using the modified-iKM algorithm.

2 Extract Coordinates from published KM curves

There are several options to extract data coordinates from published KM curves: using the Shiny app directly or other commercial digitize software such as DigitizeIt(for MAC or windows) or ScanIt(for windows).

To extract data point, first we obtain an image of the K-M graph. In general, better resolutions yield better accuracy in later steps.

We use the published locoregional control events in head and neck cancer KM curve as an example [1]. Video tutorials on data point extraction using the Shiny app and digitizeIt are are available under **User guide–video tutorials** in the Shiny app. Below, we show how ScanIt is used to do this task.
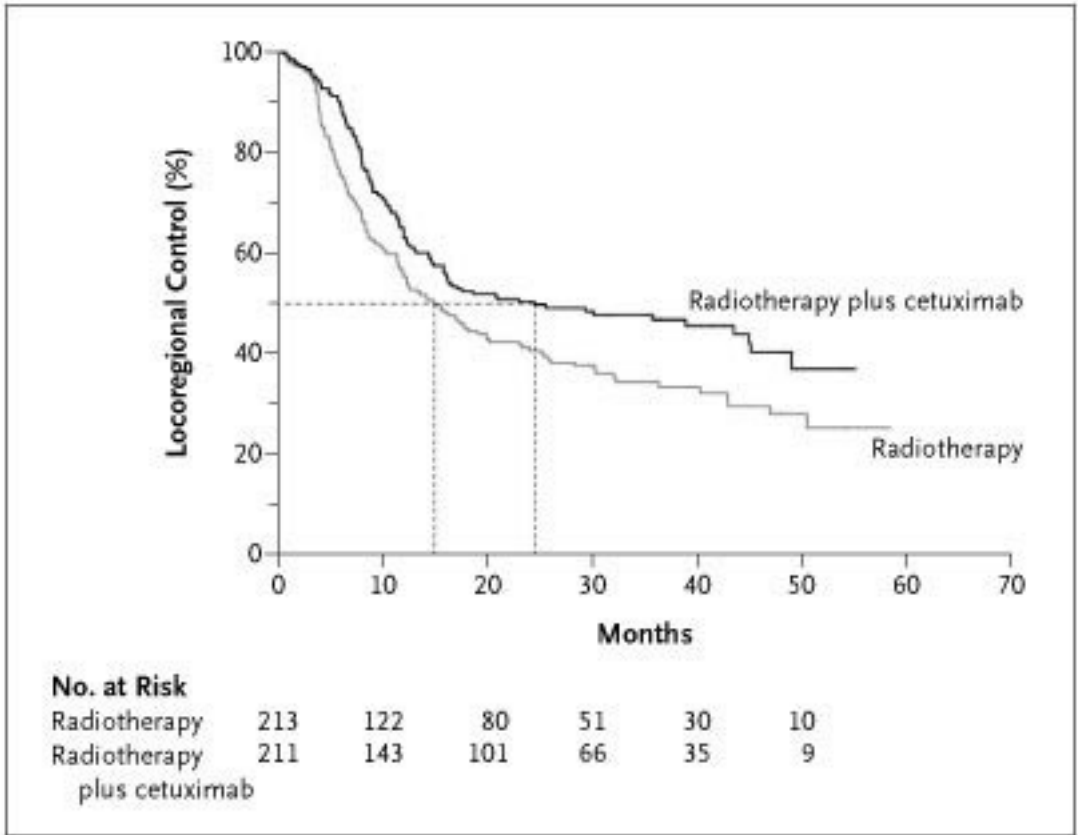
Figure 1: Locoregional control example of published Kaplan-Meier curves.

Once opening the image file, we identify the axes by clicking the intersection of the two axes and one point at each axis. A zoom view will help to pick the axis accurately (Figure 2).

Then we choose "Trace curve" mode (Figure 3), where the entire curve can be digitized by selecting only a few points (2 or more) on the curve. Since the K-M curves generally have a lot of jumps, we recommend that at least 10 to 20 points should be picked, if possible.

After tracing the K-M curve and picking some jumping points on it, the data on this curve are ready to be exported as a .txt file (Figure 4), which can be accepted by the Shiny
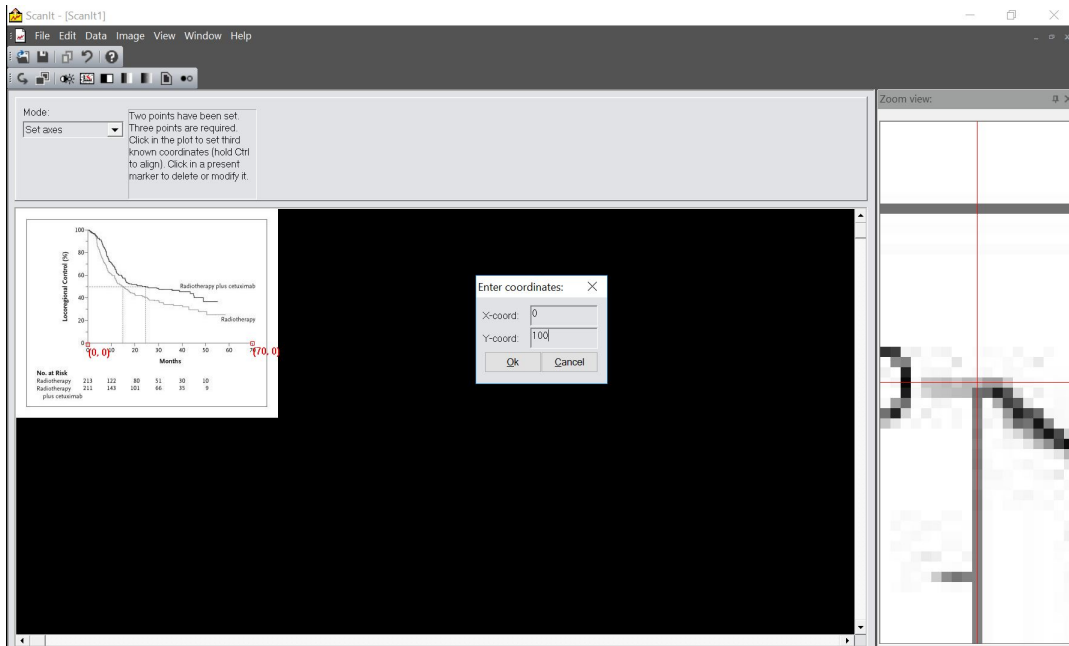
Figure 2: Setting up axis positions, under zoom view, which improves the accuracy of data extraction
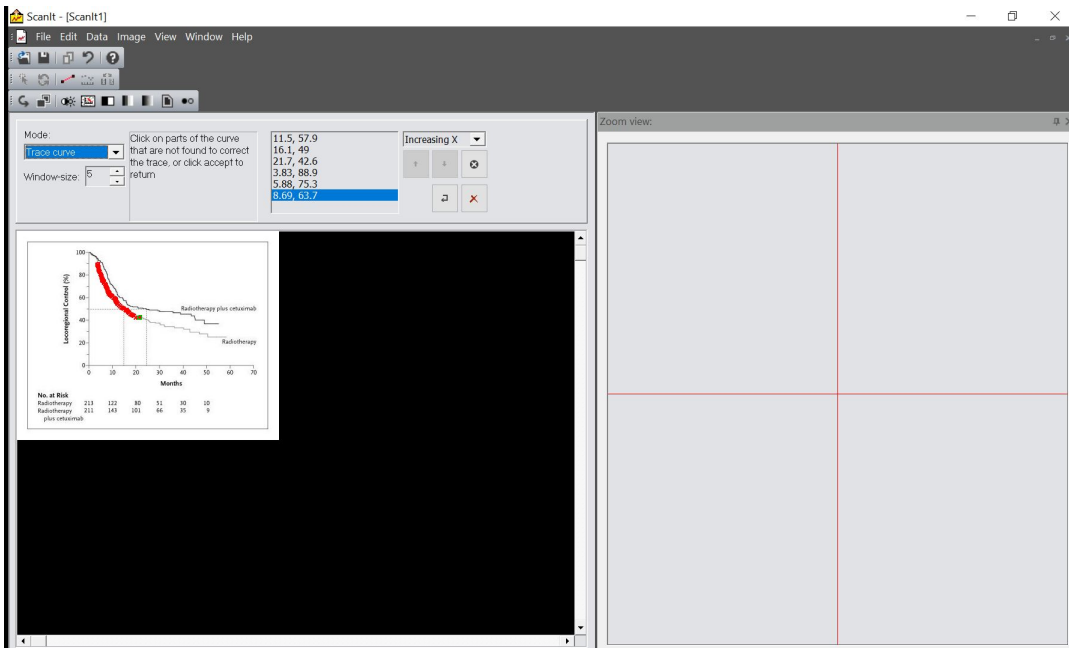

Figure 3: Using Trace curve mode to extract the entire curves automatically
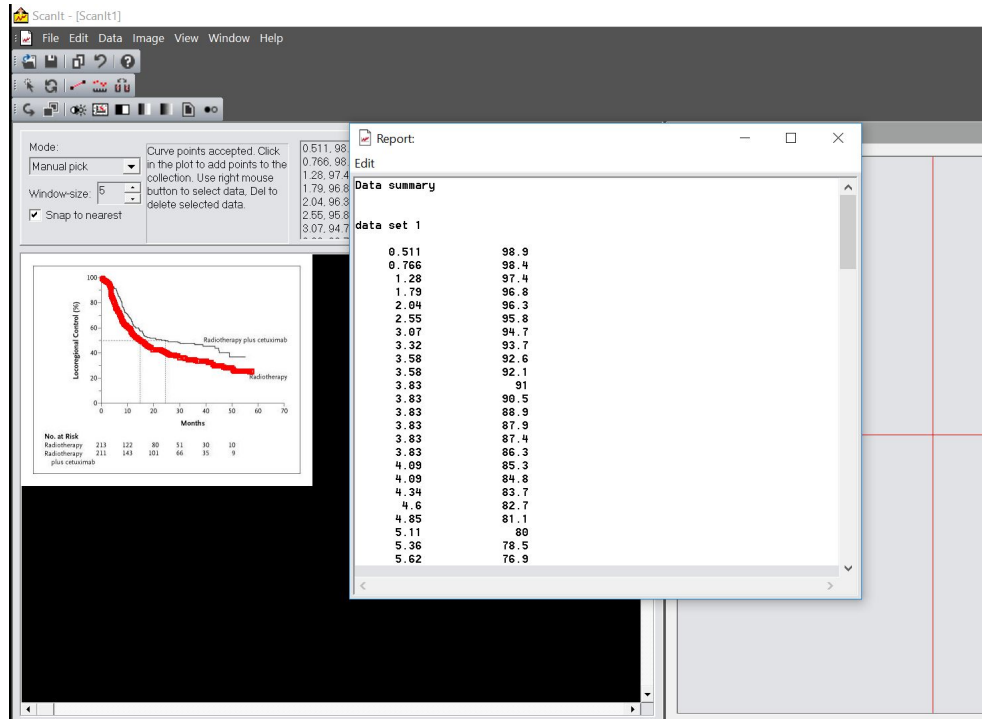We need to pick some points to let the software know where the jumps positions are.

app.

Figure 4: Data can be exported as .txt file

3 Reconstruct IPD using the Shiny app

The Shiny app has straightforward interface with the capability to

1. Extract data coordinates from published K-M curves.

2. Preprocess raw data coordinates extracted from published K-M curves.

3. Reconstruct IPD using the modified-iKM algorithm.

4. Assess the accuracy of the reconstruction.

5. Perform survival analysis using the reconstructed IPD data.

6. Download a concise report for the IPD reconstruction.

7. Provide an extensive user guide for understanding the method and using the Shiny app.

To use the Shiny app, users need the following input.

1. An image file containing K-M curve of interst and/or a .csv or .txt file containing the coordinates extracted from published K-M curves. File templates are available in the Shiny app.

2. Risk time ($trisk$).

3. Number of patients at risk ($nrisk$).

4. Total number of patients reported (not required when information for $trisk$ and $nrisk$ are provided).

5. Total number of events reported.

A video tutorial is provided under the **User guide–video tutorials–IPD reconstruction using Shiny** to demonstrate how to use the Shiny app to reconstruct IPD, evaluate the accuracy of the IPD reconstruction, and perform secondary survival analysis on reconstructed IPD, for one or two treatment groups.

References

[1] James A Bonner, Paul M Harari, Jordi Giralt, Nozar Azarnia, Dong M Shin, Roger B Cohen, Christopher U Jones, Ranjan Sur, David Raben, Jacek Jassem, et al. Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck. *New England Journal of Medicine*, 354(6):567–578, 2006.

[2] Patricia Guyot, A. E. Ades, Mario J.N.M. Ouwens, and Nicky J. Welton. Enhanced secondary analysis of survival data: Reconstructing the data from published kaplan-meier survival curves. *BMC Medical Research Methodology*, 12, 2012.