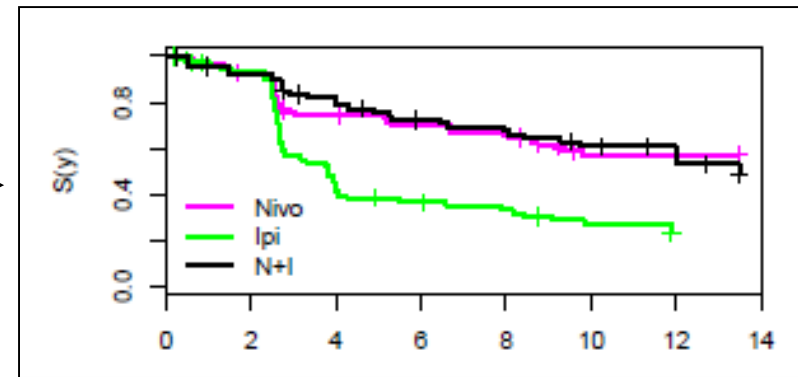
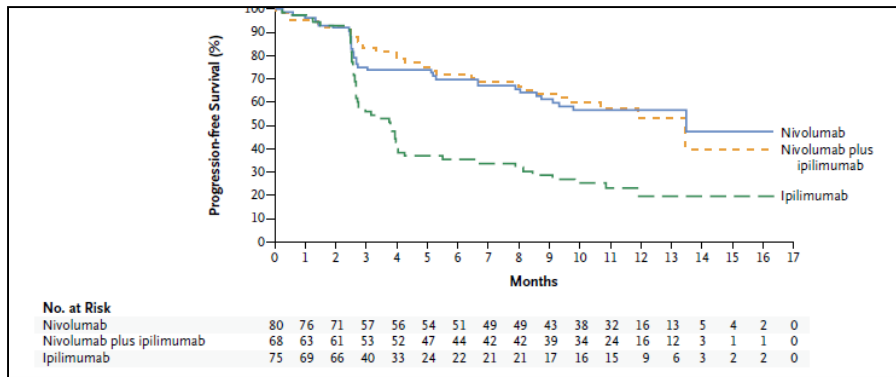


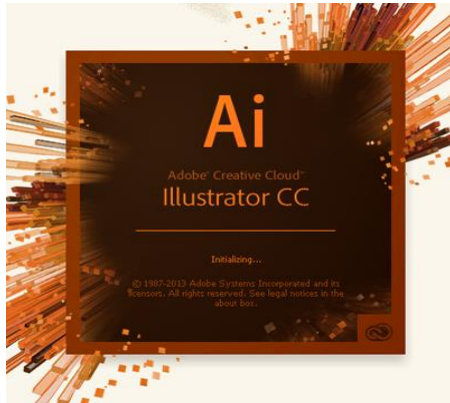
Digital Data Extraction Using R & Other Tools



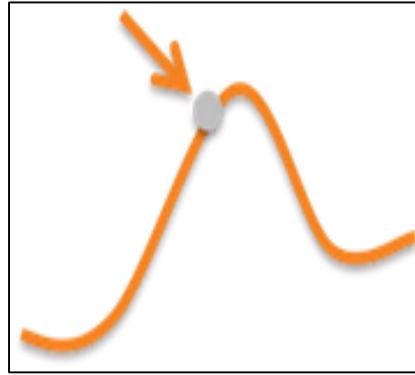
Jaya M. Satagopan
Attending Biostatistician
Memorial Sloan Kettering Cancer Center

@satagopj satagopj@mskcc.org

Combine the strengths of 3 tools and collaboration



+



+



DigitizeIt

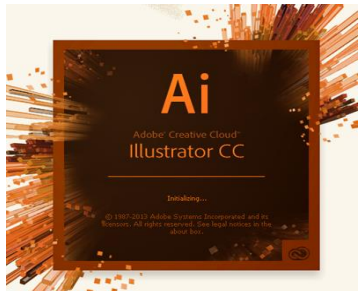
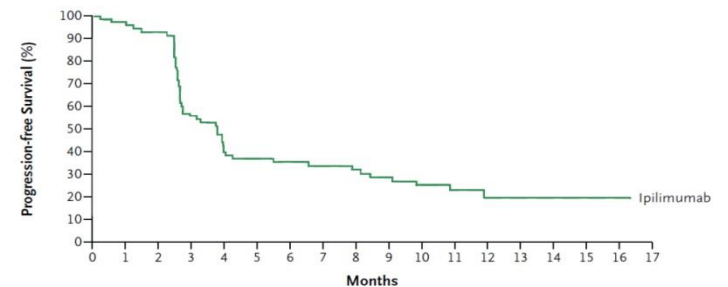
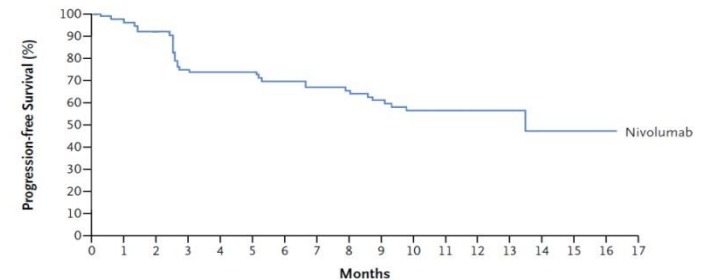
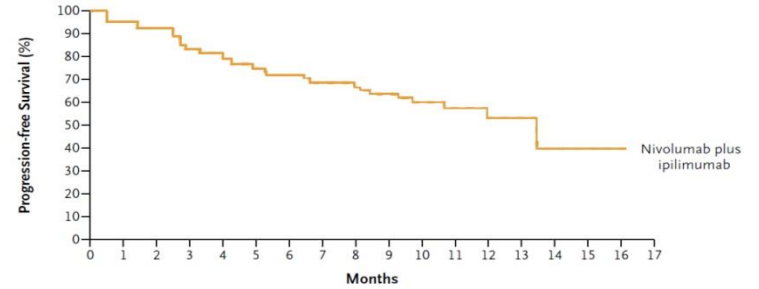
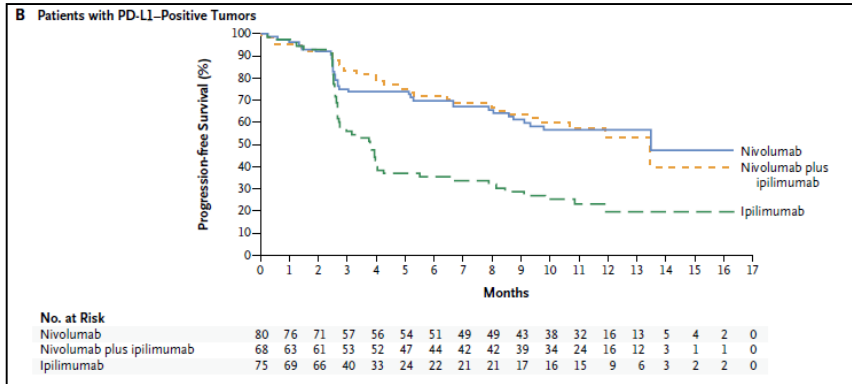


Joey Kanik



Alexia Iasonos

Step 1: Extract individual lines from published figure



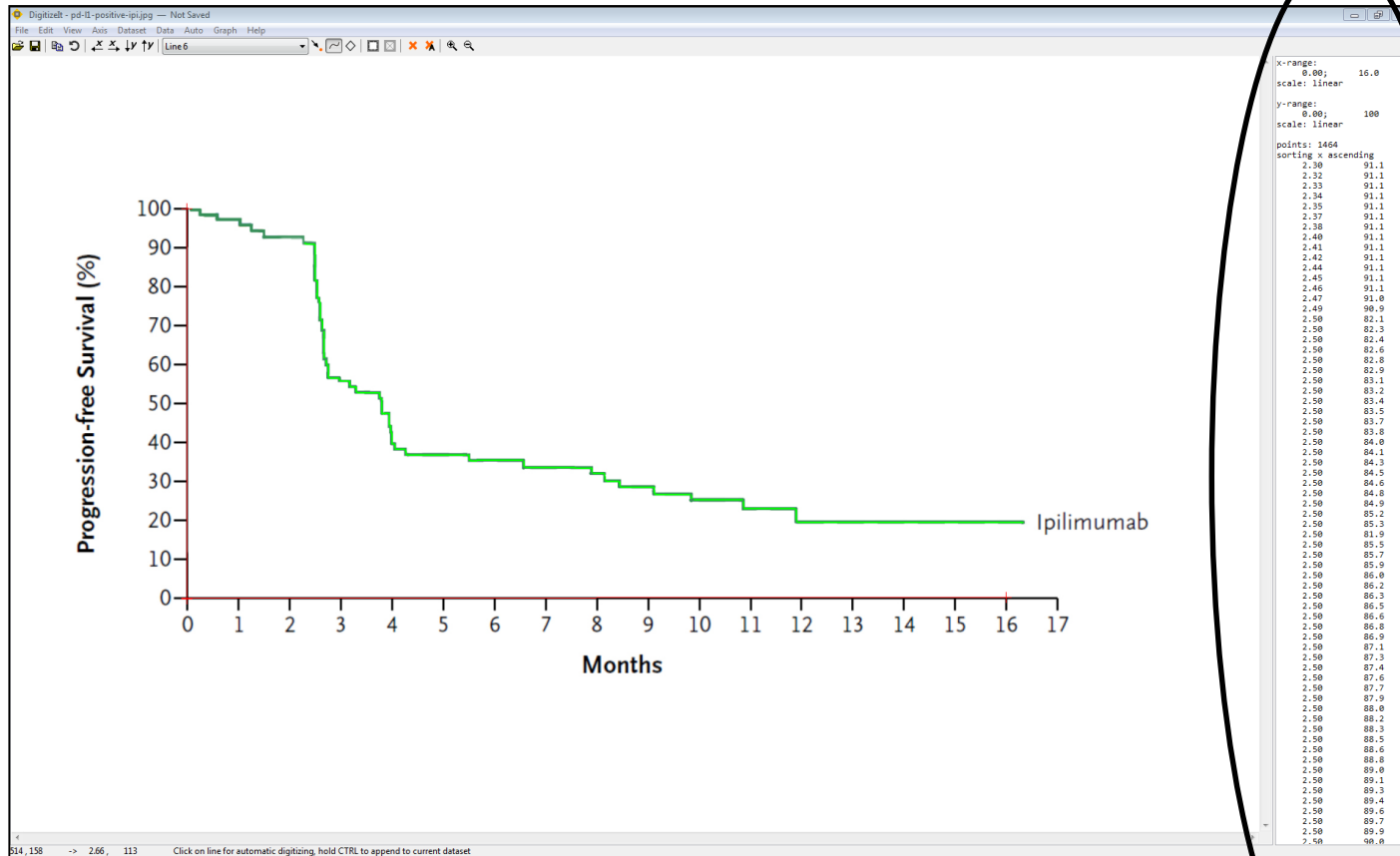
<https://www.adobe.com/products/illustrator.html>

Combined Nivolumab and Ipilimumab
or Monotherapy in Untreated Melanoma

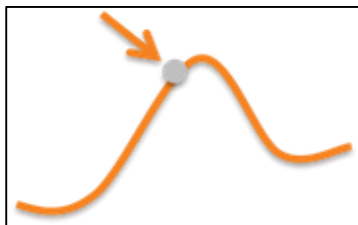
J. Larkin, V. Chiarion-Sileni, R. Gonzalez, J.J. Grob, C.L. Cowey, C.D. Lao,

N ENGL J MED 373;1 NEJM.ORG JULY 2, 2015

Step 2: Digitally extract (x,y) values



	A	B	C
1	x	y	
2		0	100
3	1.03E-01	9.95E+01	
4	1.18E-01	9.95E+01	
5	1.30E-01	9.95E+01	
6	1.41E-01	9.95E+01	
7	1.53E-01	9.95E+01	
8	1.65E-01	9.95E+01	
9	1.77E-01	9.95E+01	
10	1.89E-01	9.95E+01	
11	2.01E-01	9.95E+01	
12	2.16E-01	9.95E+01	
13	2.37E-01	9.95E+01	
14	2.47E-01	9.94E+01	
15	2.55E-01	9.93E+01	
16	2.59E-01	9.91E+01	
17	2.60E-01	9.89E+01	
18	2.61E-01	9.86E+01	
19	2.68E-01	9.85E+01	
20	2.76E-01	9.84E+01	



<https://www.digitizeit.de/>

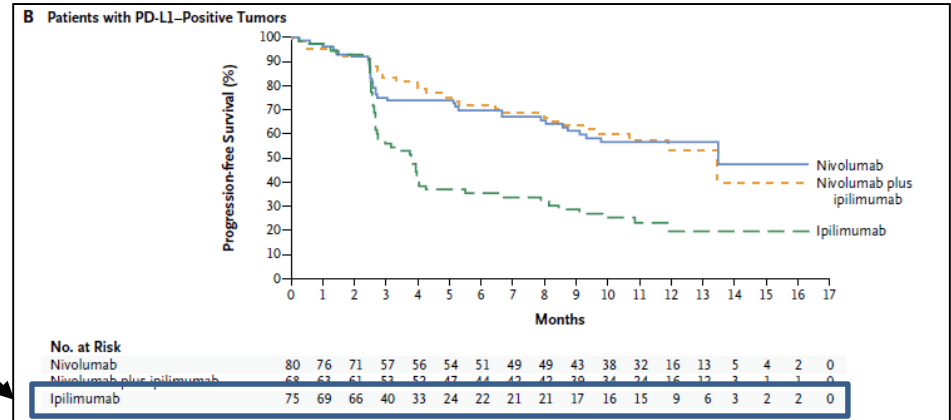
Step 3: Convert (x,y) values to person-level data

	x	y
1		
2	0	100
3	1.03E-01	9.95E+01
4	1.18E-01	9.95E+01
5	1.30E-01	9.95E+01
6	1.41E-01	9.95E+01
7	1.53E-01	9.95E+01
8	1.65E-01	9.95E+01
9	1.77E-01	9.95E+01
10	1.89E-01	9.95E+01
11	2.01E-01	9.95E+01
12	2.16E-01	9.95E+01
13	2.37E-01	9.95E+01
14	2.47E-01	9.94E+01
15	2.55E-01	9.93E+01
16	2.59E-01	9.91E+01
17	2.60E-01	9.89E+01
18	2.61E-01	9.86E+01
19	2.68E-01	9.85E+01
20	2.76E-01	9.84E+01

Pre-process



Us (pre-processing) +
Guyot et al (data extraction)



```
>
> individual.data[1:20,]
      time event tmt.arm.number treatment.type pdli.status
1 0.678     1           1      nivolumab      negative
2 0.678     1           1      nivolumab      negative
3 0.678     1           1      nivolumab      negative
4 0.678     1           1      nivolumab      negative
5 0.678     1           1      nivolumab      negative
6 0.905     1           1      nivolumab      negative
7 0.910     1           1      nivolumab      negative
8 0.939     1           1      nivolumab      negative
9 1.140     1           1      nivolumab      negative
10 1.140     1           1      nivolumab      negative
11 1.140     1           1      nivolumab      negative
12 1.390     1           1      nivolumab      negative
13 1.410     1           1      nivolumab      negative
14 1.410     1           1      nivolumab      negative
15 1.410     1           1      nivolumab      negative
16 1.610     1           1      nivolumab      negative
17 1.660     1           1      nivolumab      negative
18 1.680     1           1      nivolumab      negative
19 1.740     1           1      nivolumab      negative
20 1.910     1           1      nivolumab      negative
>
>
```

Guyot et al. *BMC Medical Research Methodology* 2012, 12:9
<http://www.biomedcentral.com/1471-2288/12/9>



TECHNICAL ADVANCE

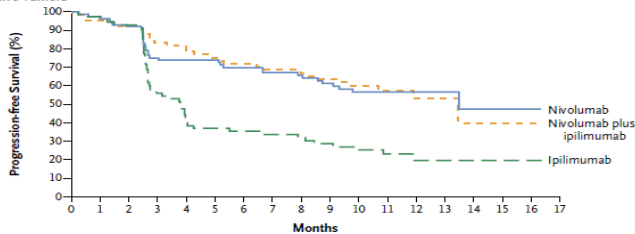
Open Access

Enhanced secondary analysis of survival data:
reconstructing the data from published
Kaplan-Meier survival curves

Patricia Guyot^{1,2*}, AE Ades¹, Mario JNM Ouwens² and Nicky J Welton¹

Digital extraction provides a good approximation

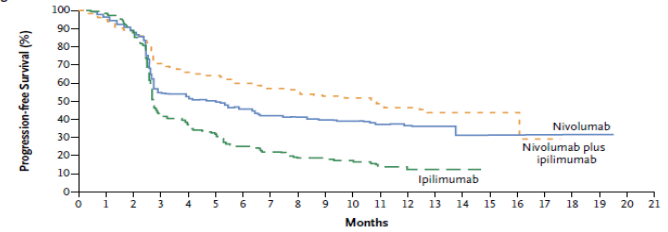
B Patients with PD-L1-Positive Tumors



No. at Risk

Nivolumab	80	76	71	57	56	54	51	49	49	43	38	32	16	13	5	4	2	0
Nivolumab plus ipilimumab	68	63	61	53	52	47	44	42	42	39	34	24	16	12	3	1	1	0
Ipilimumab	75	69	66	40	33	24	22	21	21	17	16	15	9	6	3	2	2	0

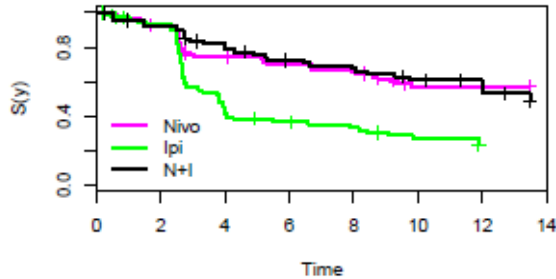
C Patients with PD-L1-Negative Tumors



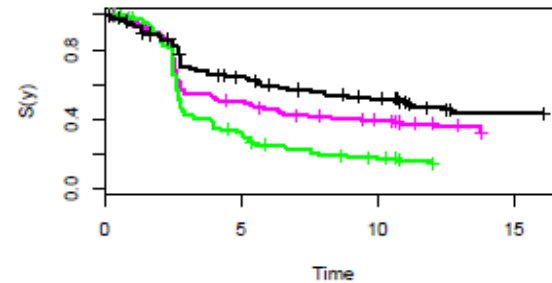
No. at Risk

Nivolumab	208	192	178	108	105	98	88	80	76	74	63	50	31	24	9	5	4	2	1	1	0
Nivolumab plus ipilimumab	210	195	181	142	134	123	112	106	105	96	88	79	42	36	13	9	6	2	1	0	0
Ipilimumab	202	183	166	82	72	59	44	39	35	31	26	22	12	8	3	1	0	0	0	0	0

Biomarker +



Biomarker -



VOLUME 34 · NUMBER 15 · MAY 20, 2016

JOURNAL OF CLINICAL ONCOLOGY

Ludovic Trinquart, Justine Jacot, Sarah C. Conner, and Raphaël Pordier

BIOMETRICS 72, 215-221
March 2016

Lihui Zhao,¹ Brian Claggett,² Lu Tian,³

VOLUME 32 · NUMBER 22 · AUGUST 1 2014

JOURNAL OF CLINICAL ONCOLOGY

Hajime Uno, Brian Claggett, Lu Tian, 1

Clinical data sharing is a vast and heavily discussed topic



Our effort: a temporary solution for an immediate need using Adobe I, DigitizeIt & R



Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib



Data Article

A reconstructed melanoma data set for
evaluating differential treatment benefit
according to biomarker subgroups



Jaya M. Satagopan^{*}, Alexia Iasonos, Joseph G. Kanik

Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA

For R codes and digitized data:

<https://www.mskcc.org/departments/epidemiology-biostatistics/biostatistics/data-brief-full-data-set>