HOW DO TYPICAL RUNNERS' PERFORMANCES VARY WITH AGE AND GENDER?

Richard L Smith and Spencer Siegel (University of North Carolina); Dorit Hammerling (Colorado School of Mines)

New England Symposium on Statistics in Sports

September 28, 2019

samsi
NSF Duke NCSU UNC

UNC
COLLEGE OF
ARTS & SCIENCES

BOSTON ATHLETIC ASSOCIATION
BOSTON MARATHON

Acknowledgements

- The staff of the Boston Athletic Association – Jack Fleming (COO), Tom Grilk (CEO), Michael Pieroni, Chris Menard
- Scott Powers and Jessi Cisewski-Kehe were collaborators in the original project (see paper in *Chance*, 2014)
- The work I'm presenting here uses data originally collected for a different study of Boston Marathon performances (Hammerling et al., *PLoS ONE*, 2014)
- Additional work by Spencer (senior thesis, in progress)
- Laura Albrecht, Ross Ring-Jarvi and Dorit Hammerling have a poster on a related topic. They are focusing on a different aspect of the same set of questions.

this shoe represents the millions of people across the world who support the boston marathon and the entire running community.
this year, no matter where you are, we all run boston.
energy boost 2.0 limited edition, 2014
#weallrunboston
Boston Athletic Association
COMMEMORATIVE PROGRAM
26 MILES TO BOSTON
The Boston Marathon Experience from Hopkinton to Copley Square
Racers' Record Book

2020 Boston Marathon qualifying

Qualifying standard and actual qualification time by age group and gender

Age group	men's standard	men's qualification	women's standard	women's qualification
18-34	3:00:00	2:58:21	3:30:00	3:28:21
35-39	3:05:00	3:03:21	3:35:00	3:33:21
40-44	3:10:00	3:08:21	3:40:00	3:38:21
45-49	3:20:00	3:18:21	3:50:00	3:48:21
50-54	3:25:00	3:23:21	3:55:00	3:53:21
55-59	3:35:00	3:33:21	4:05:00	4:03:21
60-64	3:50:00	3:48:21	4:20:00	4:18:21
65-69	4:05:00	4:03:21	4:35:00	4:33:21
70-74	4:20:00	4:18:21	4:50:00	4:48:21
75-79	4:35:00	4:33:21	5:05:00	5:03:21
80 and older	4:50:00	4:48:21	5:20:00	5:18:21

Source: BAA (reprinted by the Boston Globe)

YEAR	FIELD SIZE	"CUT-OFF TIME"*	QUALIFIERS NOT ACCEPTED
2012	27,000	1:14	3,228
2014	36,000	1:38	2,976
2015	30,000	1:02	1,947
2016	30,000	2:28	4,562
2017	30,000	2:09	2,957
2018	30,000	3:23	5,062
2019	30,000	4:52	7,248
2020	31,500	1:39	3,161

Source: BAA

BACKGROUND

- The Boston Marathon is the only major marathon to require qualifying standards of the majority of participants
- Qualifying standards were first introduced in the 1970s and have been revised several times since
- I first got involved in 2010, after entries for the 2011 race closed out in one day.
- With Scott and Jessi, I wrote a report that provided some background for the 2012 revision of the standards
- Since 2012, the race has continued to grow and this year the standards were revised again

Table 1: Boston Marathon Qualifying Times (up to 2012)

Age	Men	Women
18-34	3hrs 10min	3hrs 40min
35-39	3hrs 15min	3hrs 45min
40-44	3hrs 20min	3hrs 50min
45-49	3hrs 30min	4hrs 00min
50-54	3hrs 35min	4hrs 05min
55-59	3hrs 45min	4hrs 15min
60-64	4hrs 00min	4hrs 30min
65-69	4hrs 15min	4hrs 45min
70-74	4hrs 30min	5hrs 00min
75-79	4hrs 45min	5hrs 15min
80 and over	5hrs 00min	5hrs 30min

CHANGES SINCE 2012

- In 2012, the standards were uniformly tightened by five minutes (e.g. 3 hours, 5 minutes for men 18-34)
- Beginning in 2020, the standards have been further tightened by five minutes (e.g. 3 hours, 0 minutes for men 18-34)
- On both occasions, I have made recommendations about how these changes are likely to affect the size of the field
- However, the basis of the standards themselves has not been changed (e.g. 30-minute gap between qualifying standards for men and women in the same age group)
- The current work is motivated by broader questions of equity across different age and gender categories

IDEA MOTIVATING THE PRESENT TALK

- All runners slow down as they get older, but there is a lot of individual variability
- Try to use statistical methods to characterize the age-graded performance of a “typical” runner
- The standard method used for age-graded performances is nominally based on world records in different age groups, but this may not reflect typical runners’ performances
- The approach I am presenting here does not directly address the gender-equity issue (see poster by Albrecht et al.) though the two questions are closely related

Longitudinal Approach to the Performance v. Age Problem

- About 500 runners have run the Boston marathon at least 10 years in succession (BAA)

Longitudinal Approach to the Performance v. Age Problem

- About 500 runners have run the Boston marathon at least 10 years in succession (BAA)
- No easy way to reconstruct the list

Longitudinal Approach to the Performance v. Age Problem

- About 500 runners have run the Boston marathon at least 10 years in succession (BAA)
- No easy way to reconstruct the list
- I used the datasets I had available to find 1,272 runners who had run each of 2010, 2011, 2013

Longitudinal Approach to the Performance v. Age Problem

- About 500 runners have run the Boston marathon at least 10 years in succession (BAA)
- No easy way to reconstruct the list
- I used the datasets I had available to find 1,272 runners who had run each of 2010, 2011, 2013
- I then used the BAA archive to find all of those runners who had run at least 6 times (men) or 5 times (women) during 2001-2013

Longitudinal Approach to the Performance v. Age Problem

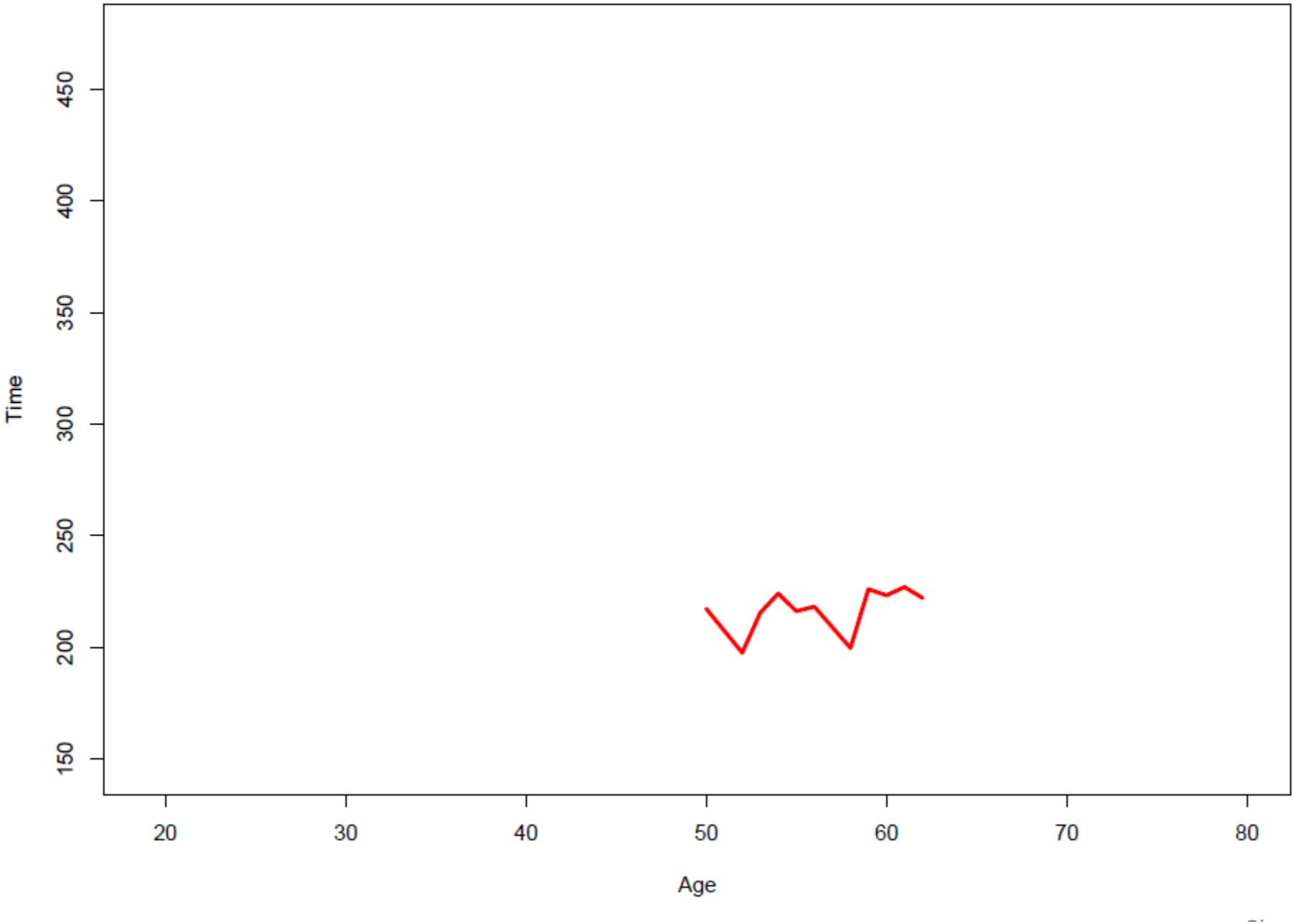
- About 500 runners have run the Boston marathon at least 10 years in succession (BAA)
- No easy way to reconstruct the list
- I used the datasets I had available to find 1,272 runners who had run each of 2010, 2011, 2013
- I then used the BAA archive to find all of those runners who had run at least 6 times (men) or 5 times (women) during 2001-2013
- Runners who did not finish in 2013 were estimated using Hammerling et al. (2014)

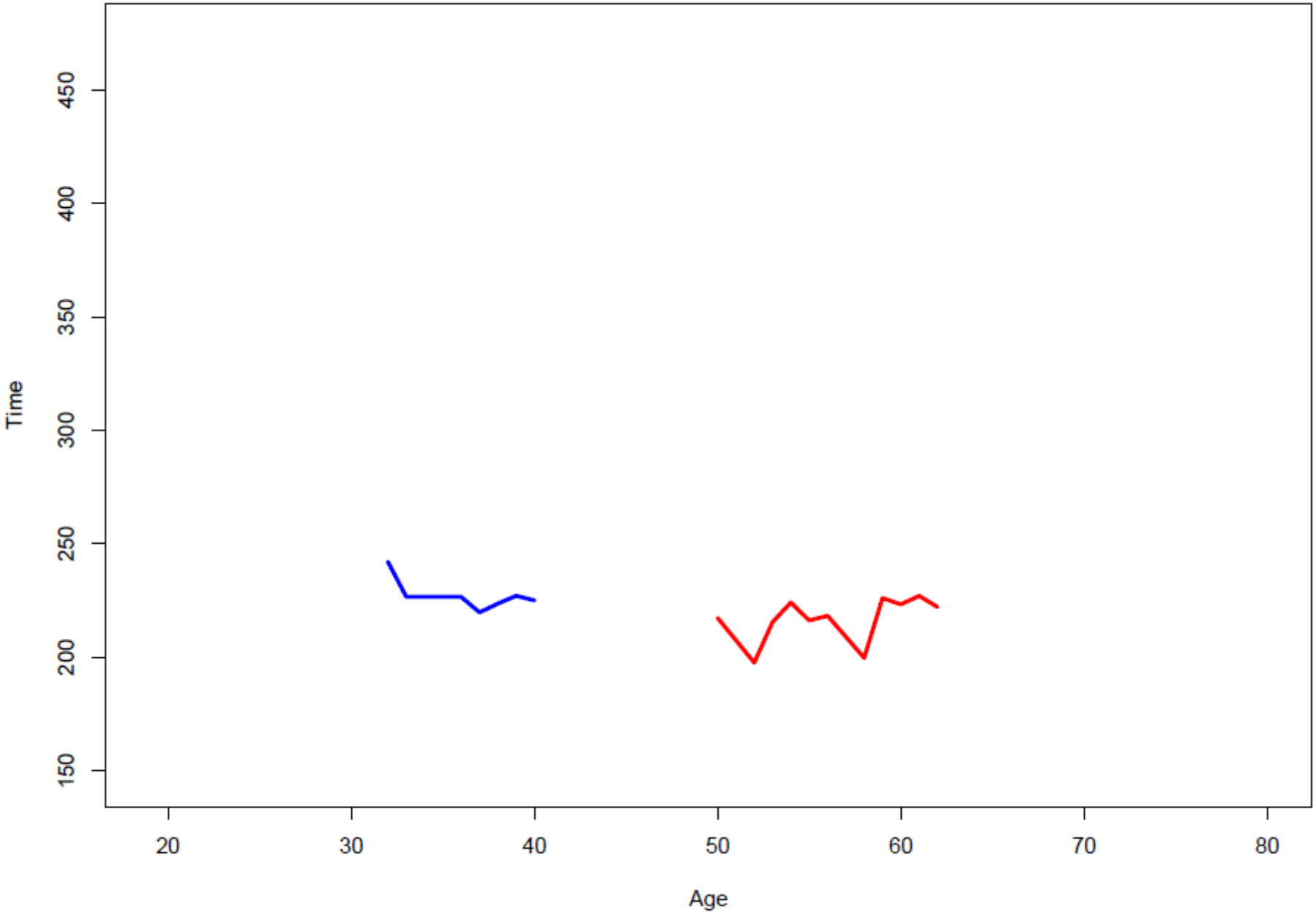
Longitudinal Approach to the Performance v. Age Problem

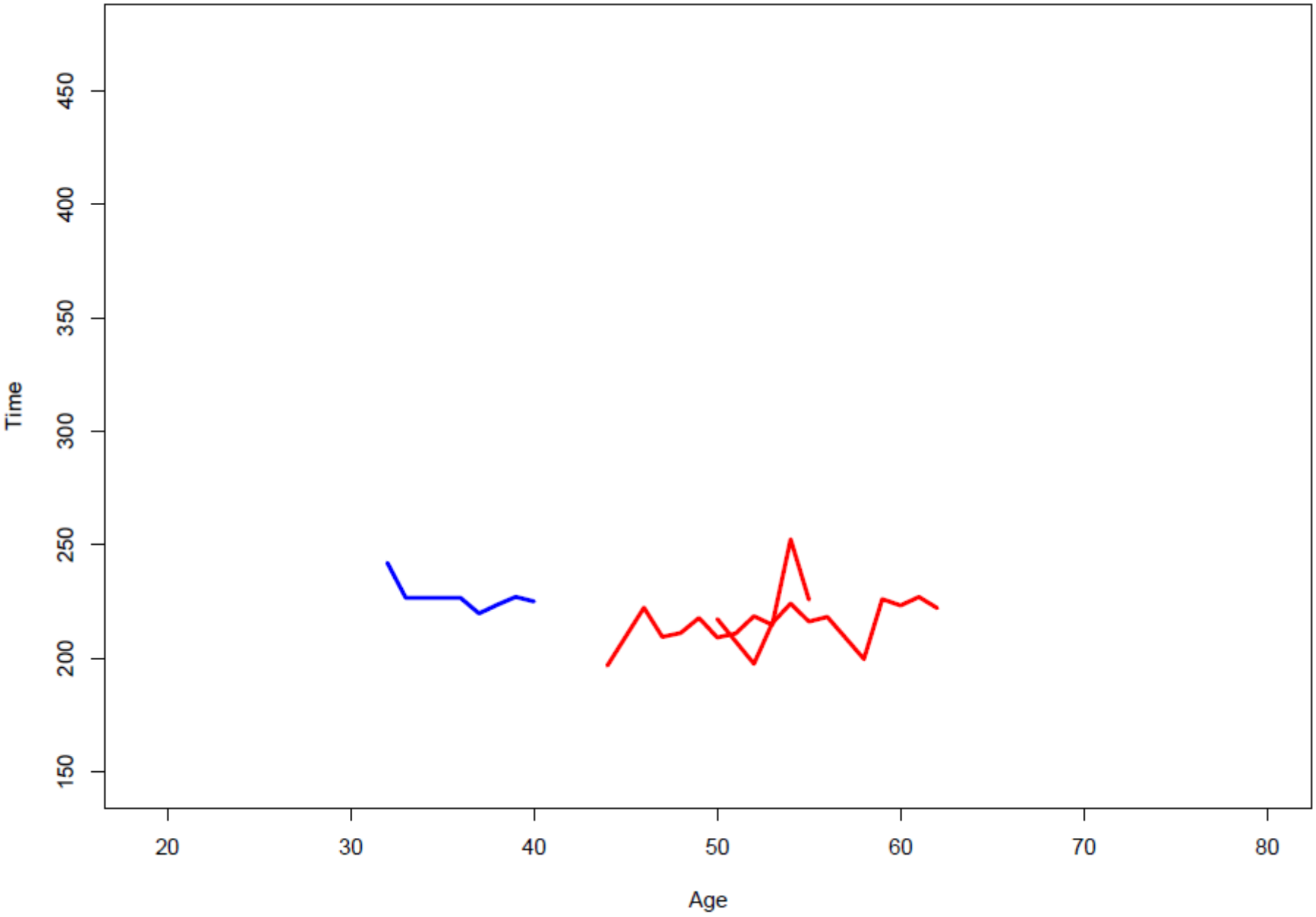
- About 500 runners have run the Boston marathon at least 10 years in succession (BAA)
- No easy way to reconstruct the list
- I used the datasets I had available to find 1,272 runners who had run each of 2010, 2011, 2013
- I then used the BAA archive to find all of those runners who had run at least 6 times (men) or 5 times (women) during 2001-2013
- Runners who did not finish in 2013 were estimated using Hammerling et al. (2014)
- Result: 547 men and 249 women identified (806 runners; 7,219 individual race results)

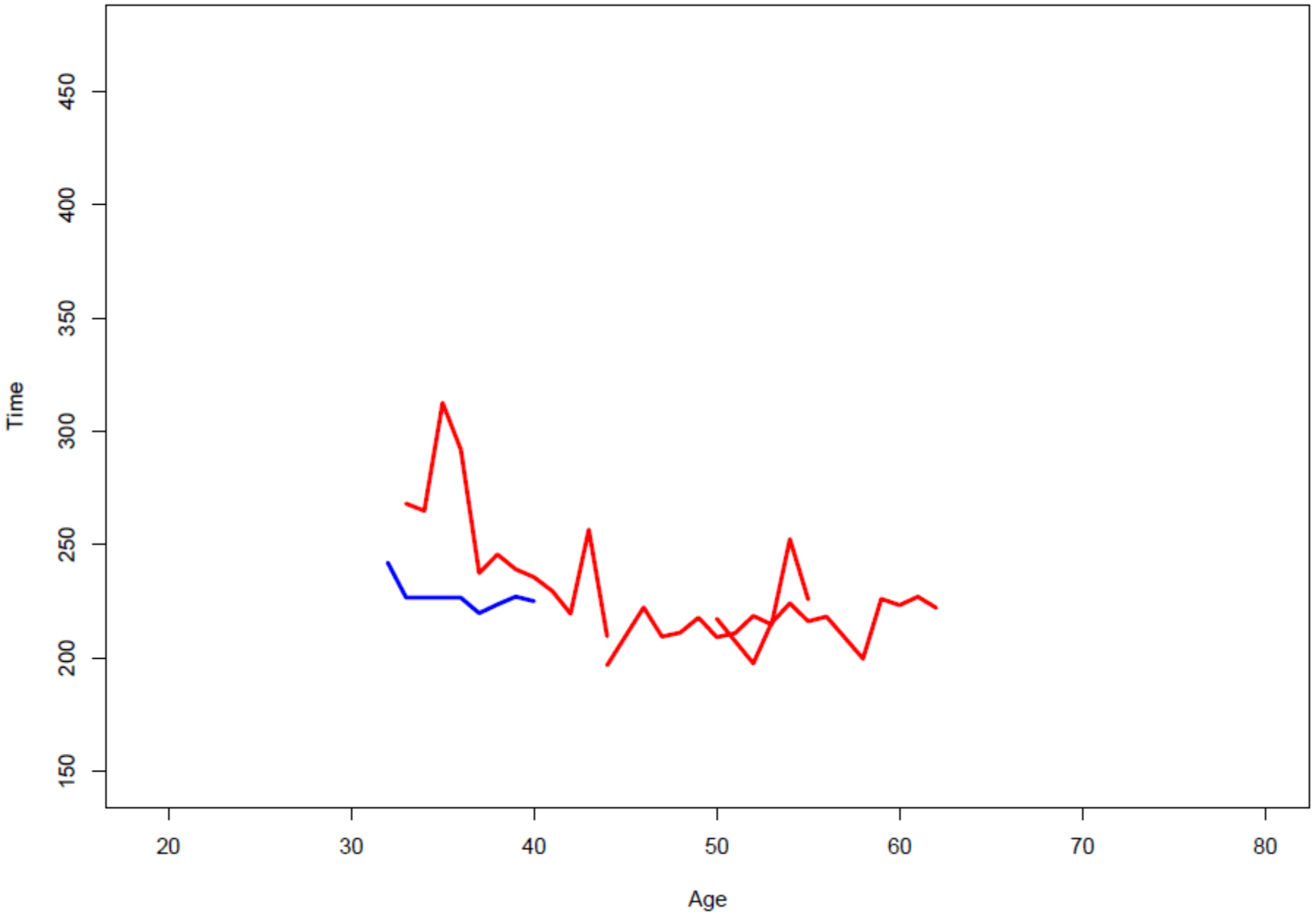
Idea of the Approach

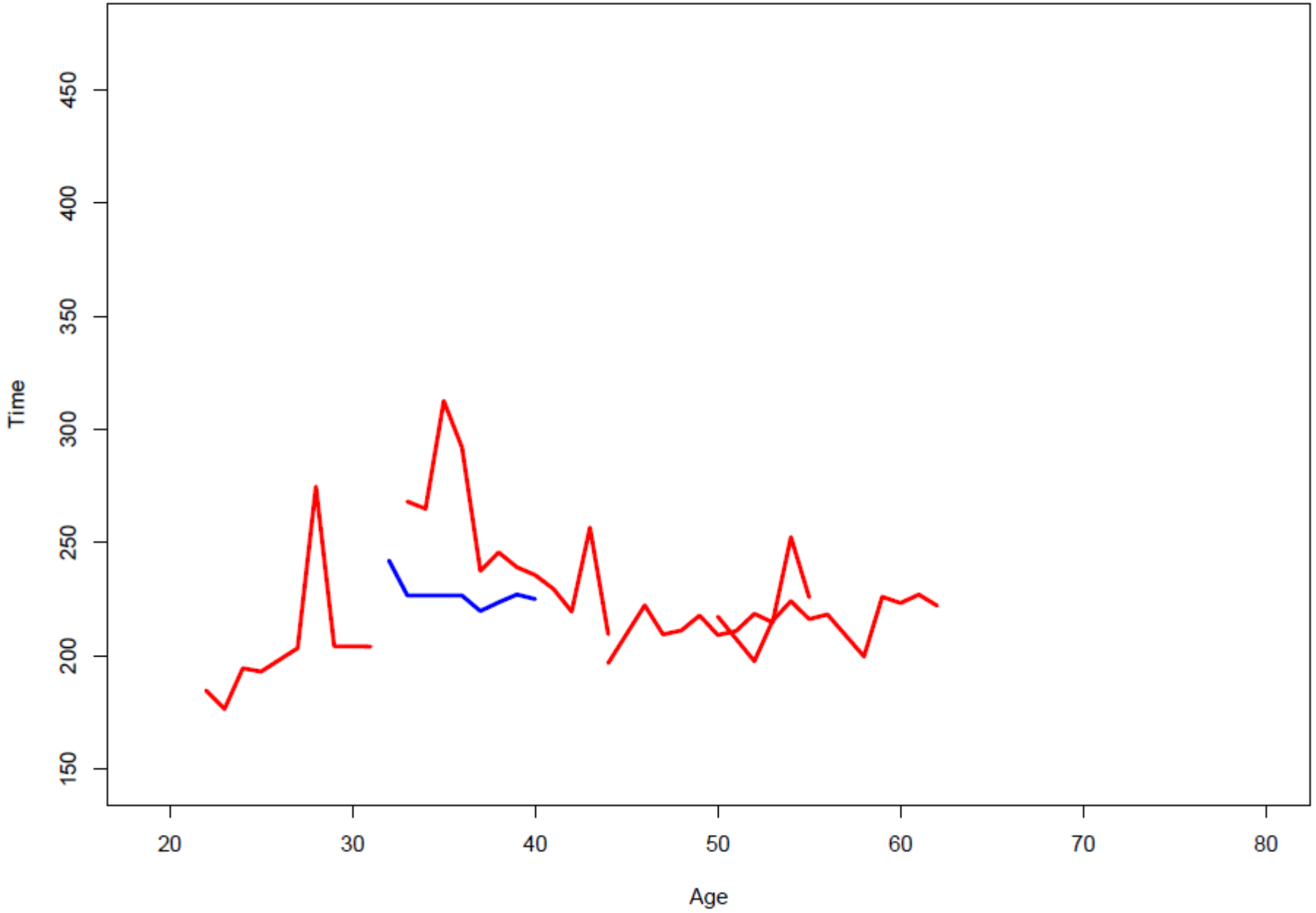
- Each individual runner defines their own age v. race time plot
- Color code to distinguish men (red) from women (blue)
- How best to combine information across runners?

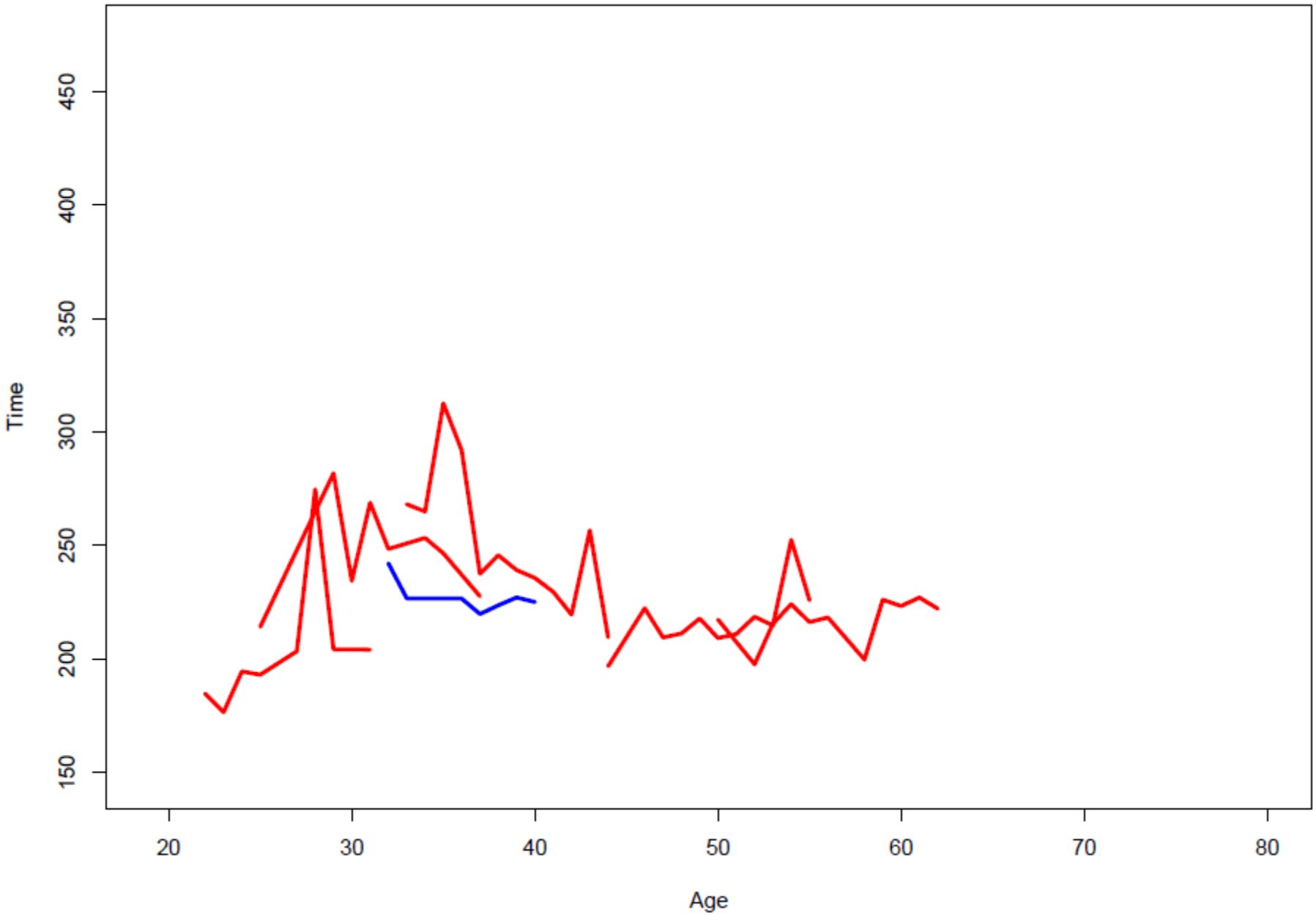

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

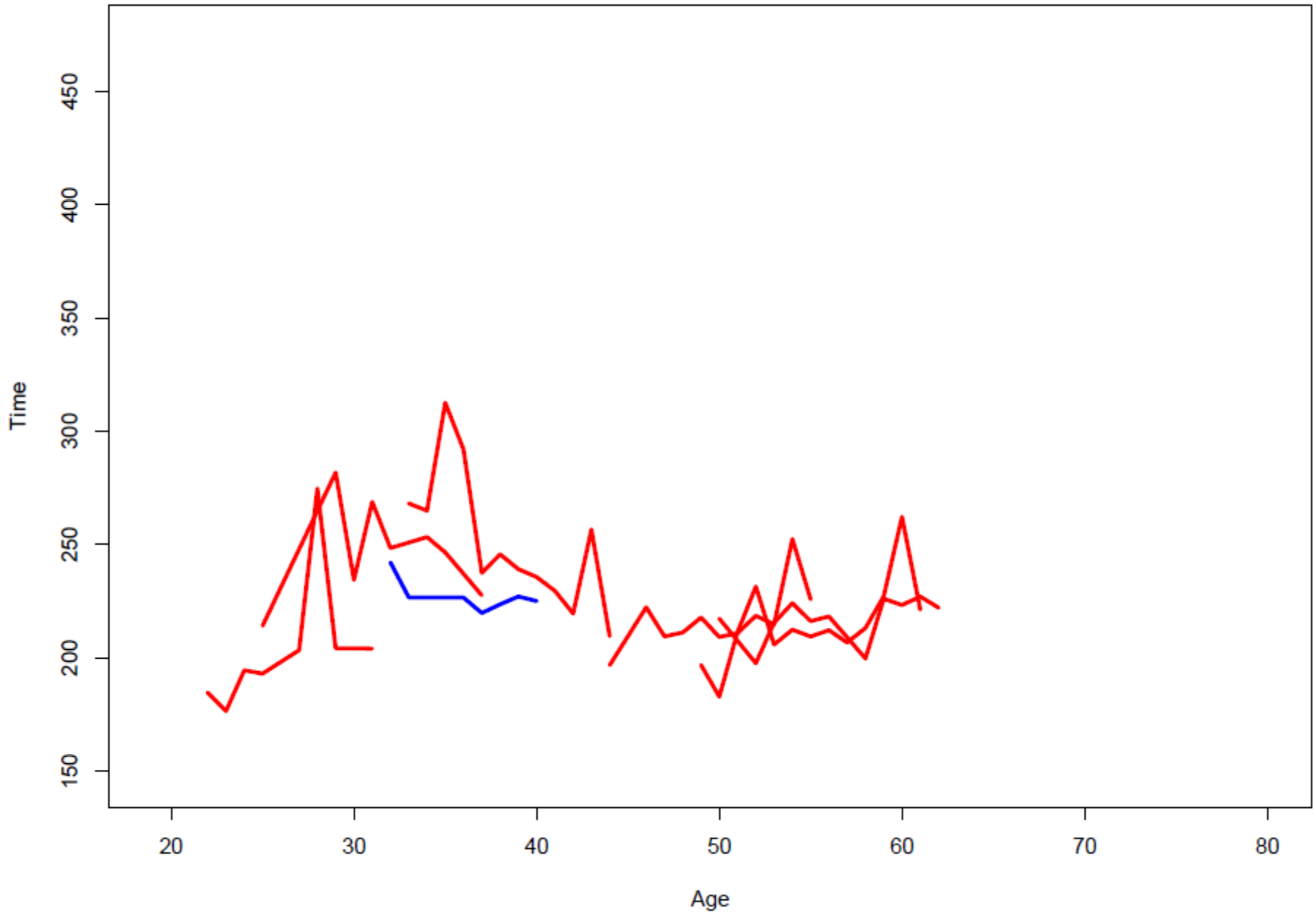

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

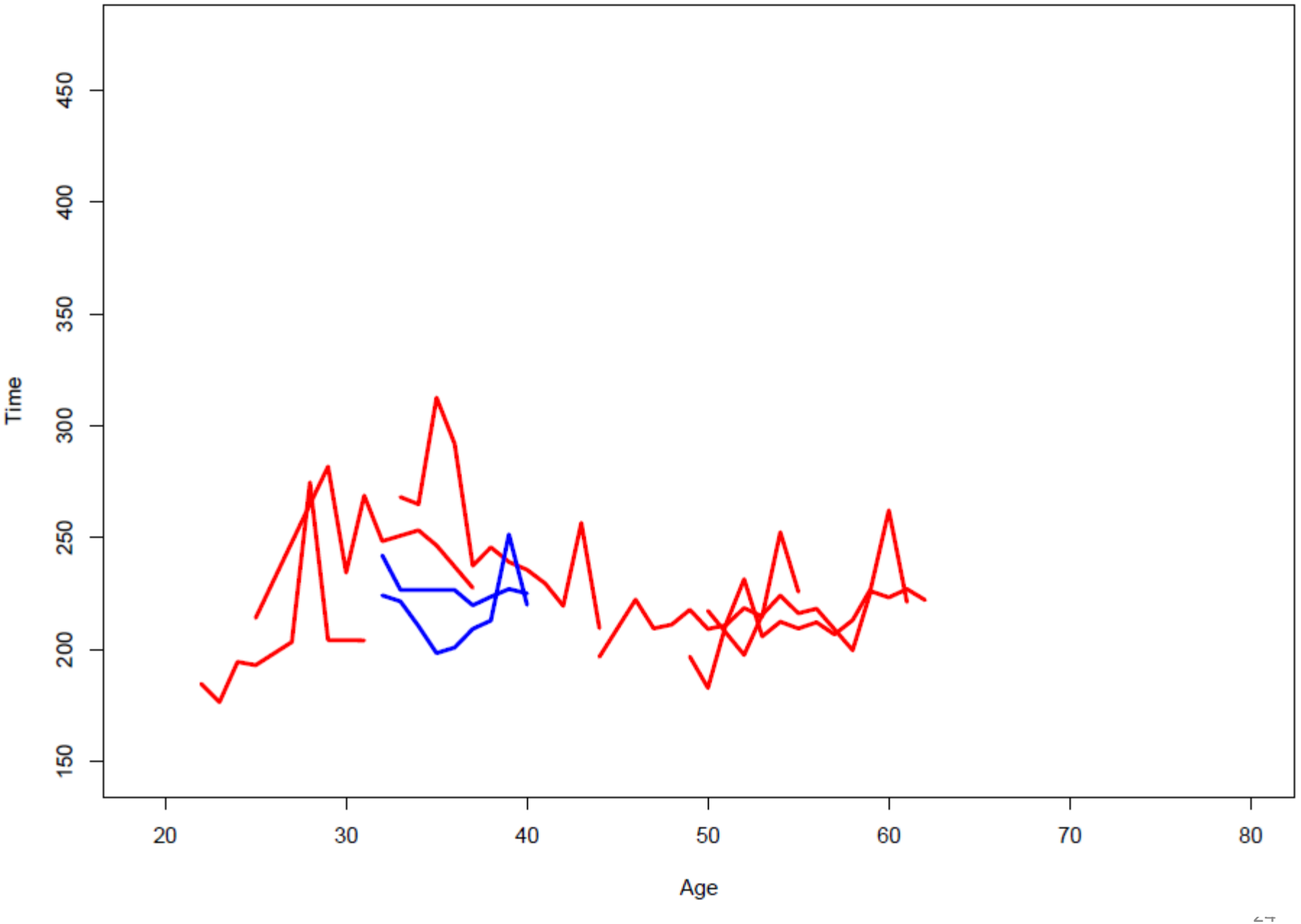

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

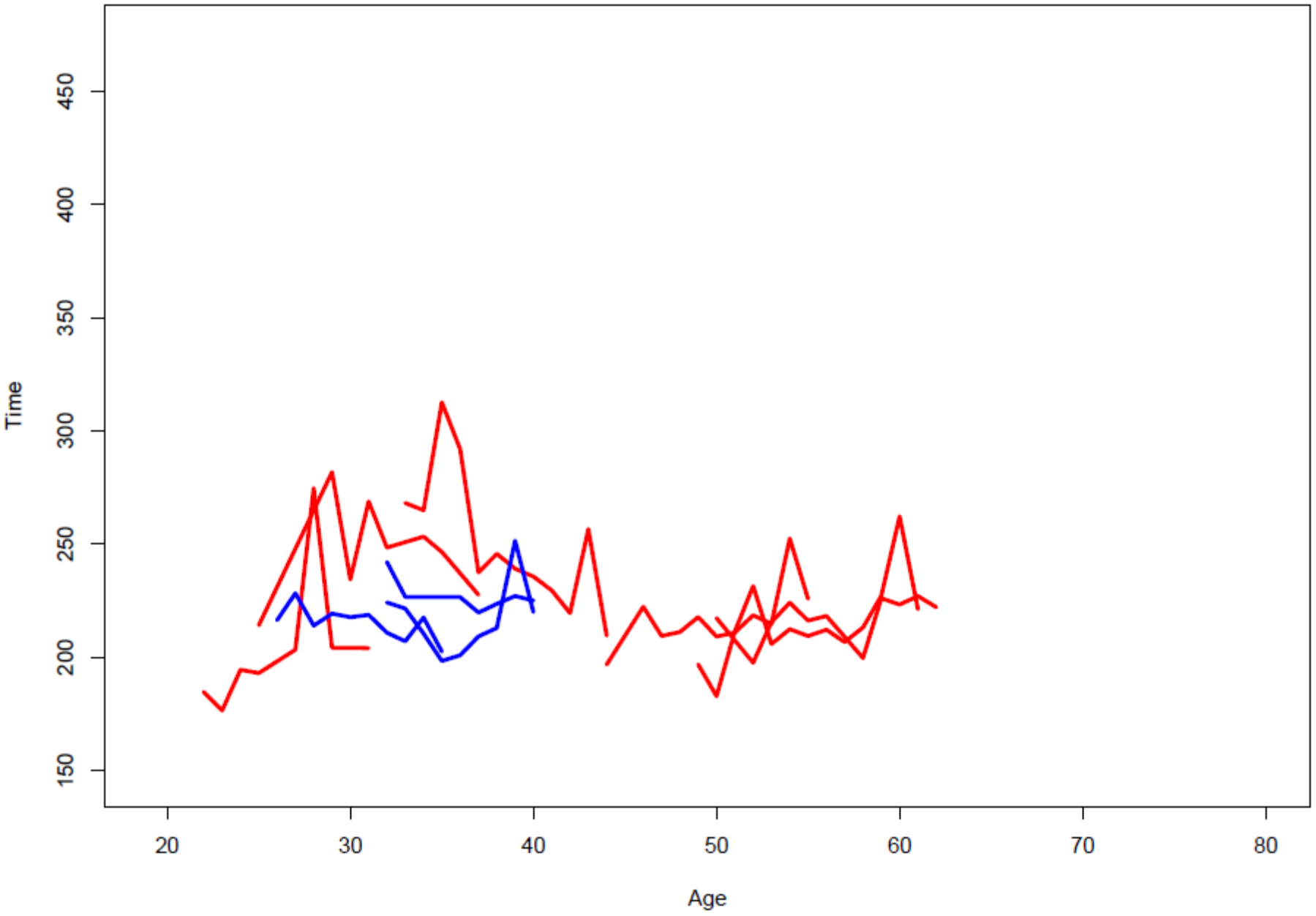

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

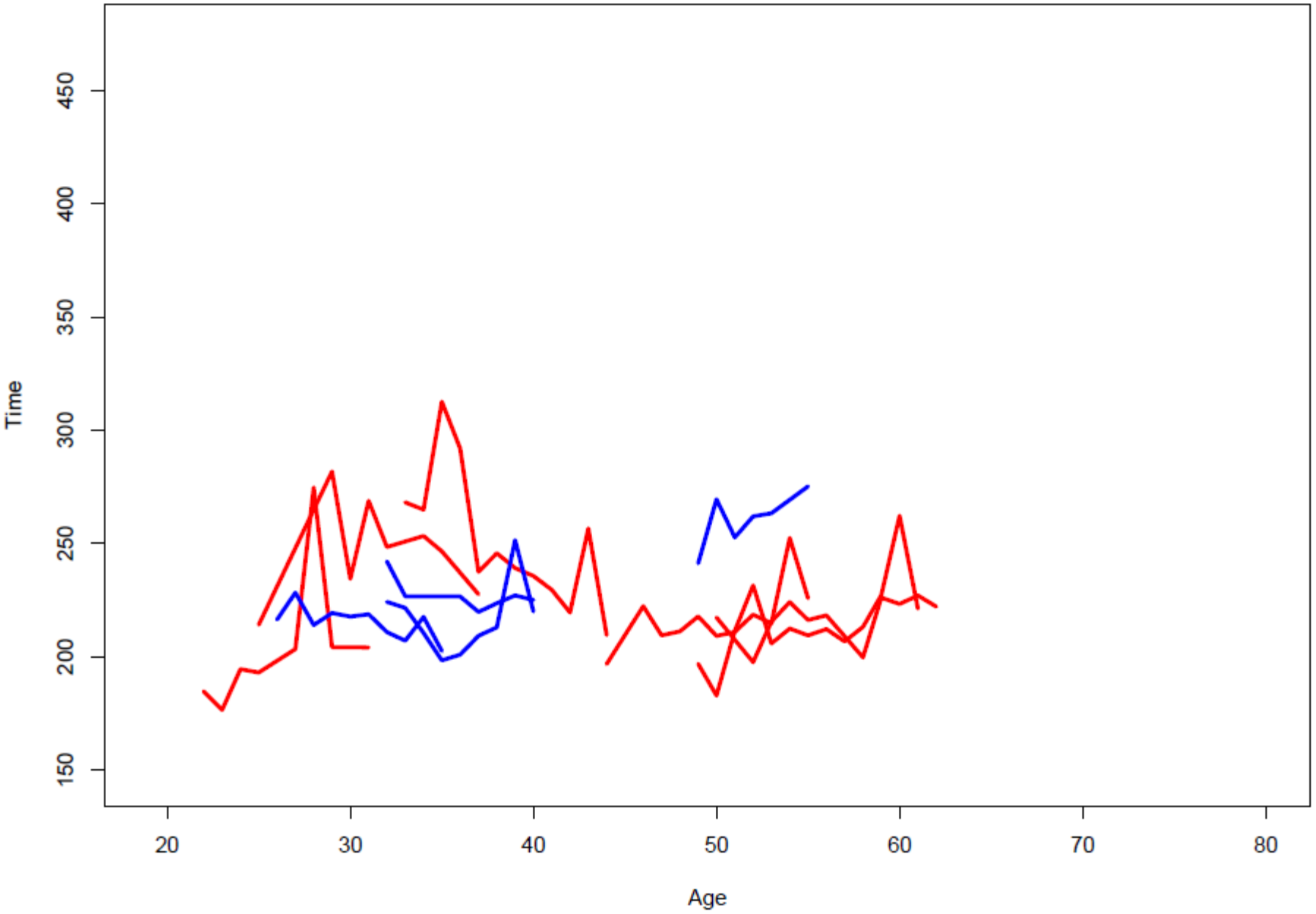

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

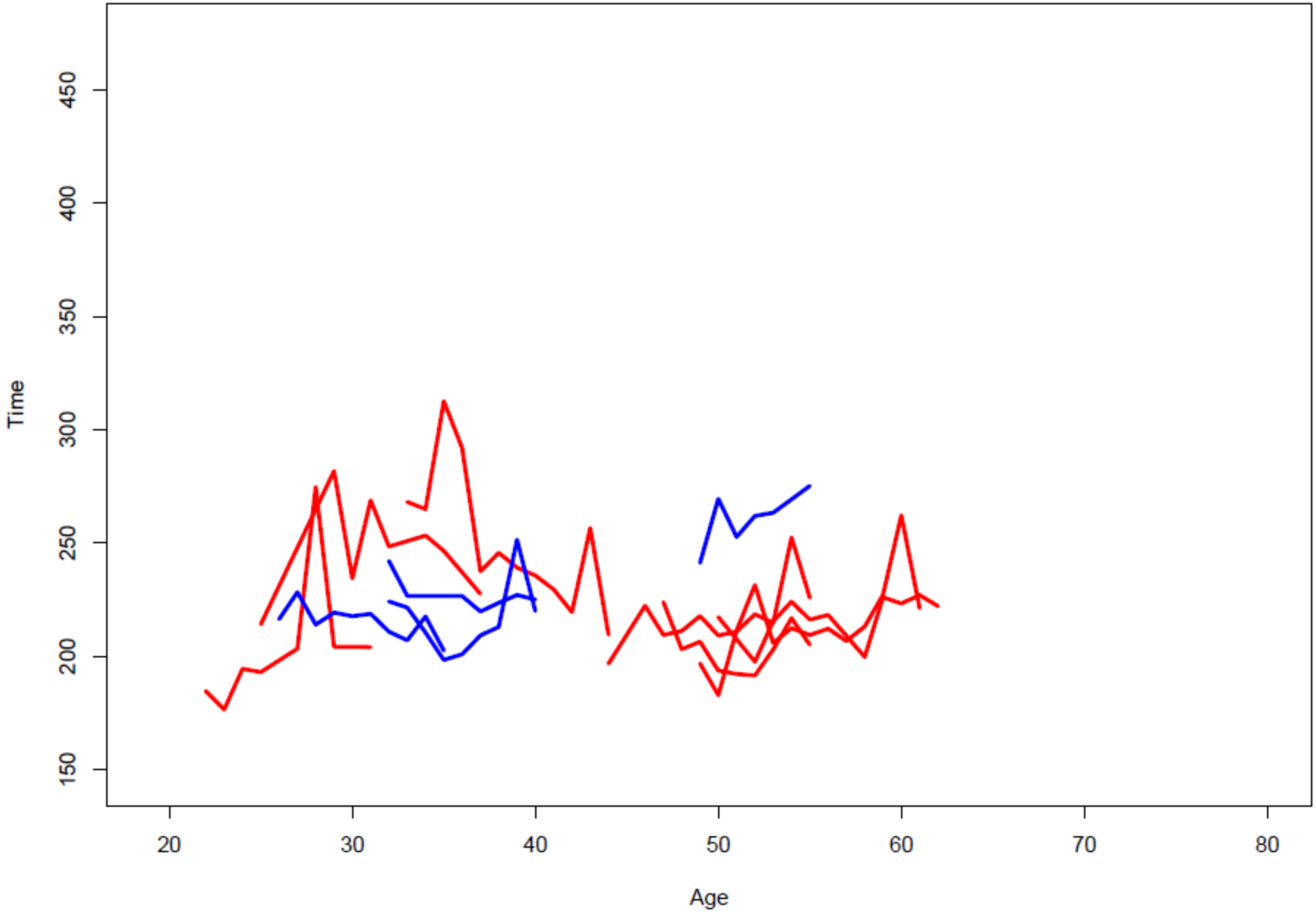

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

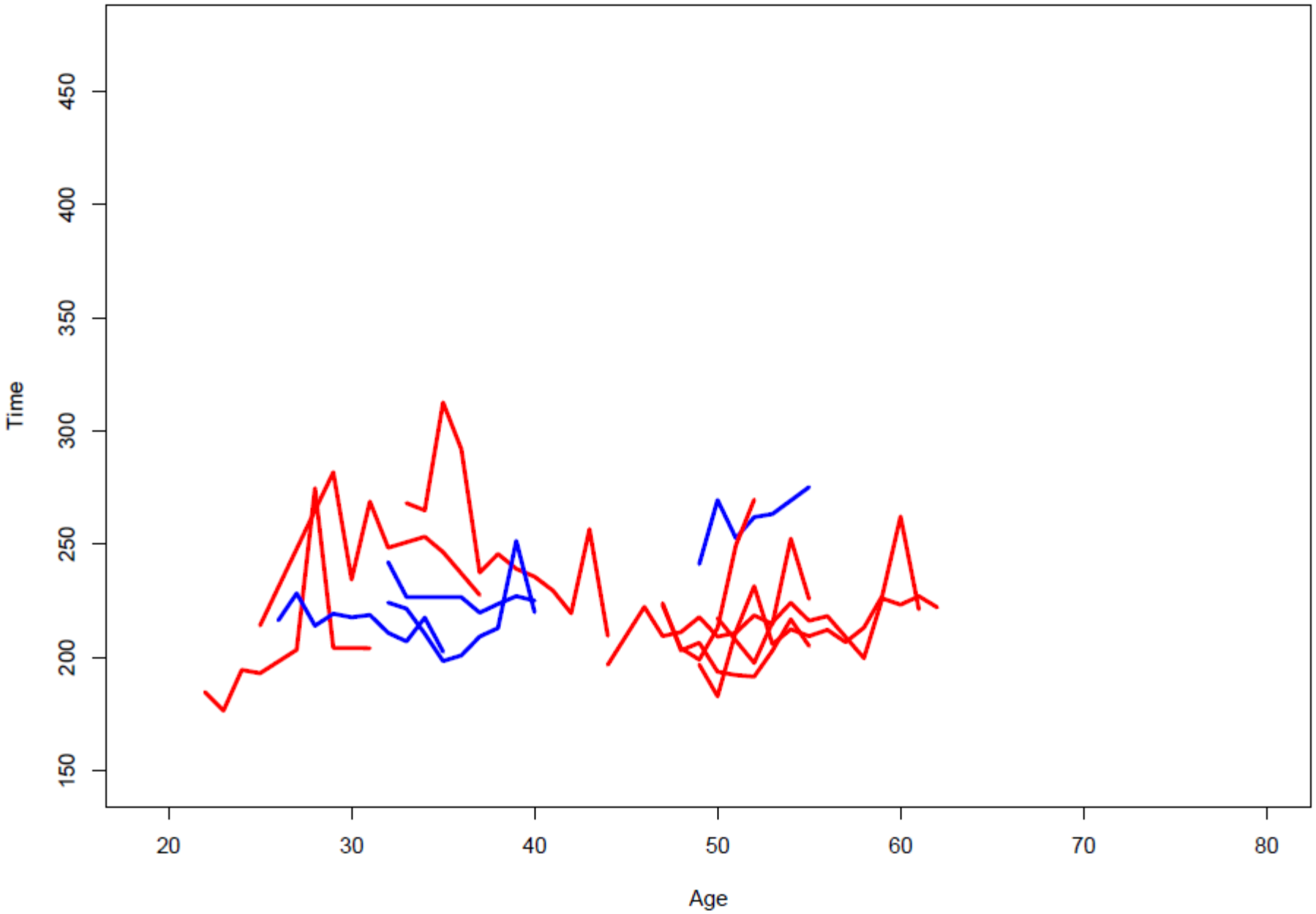

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

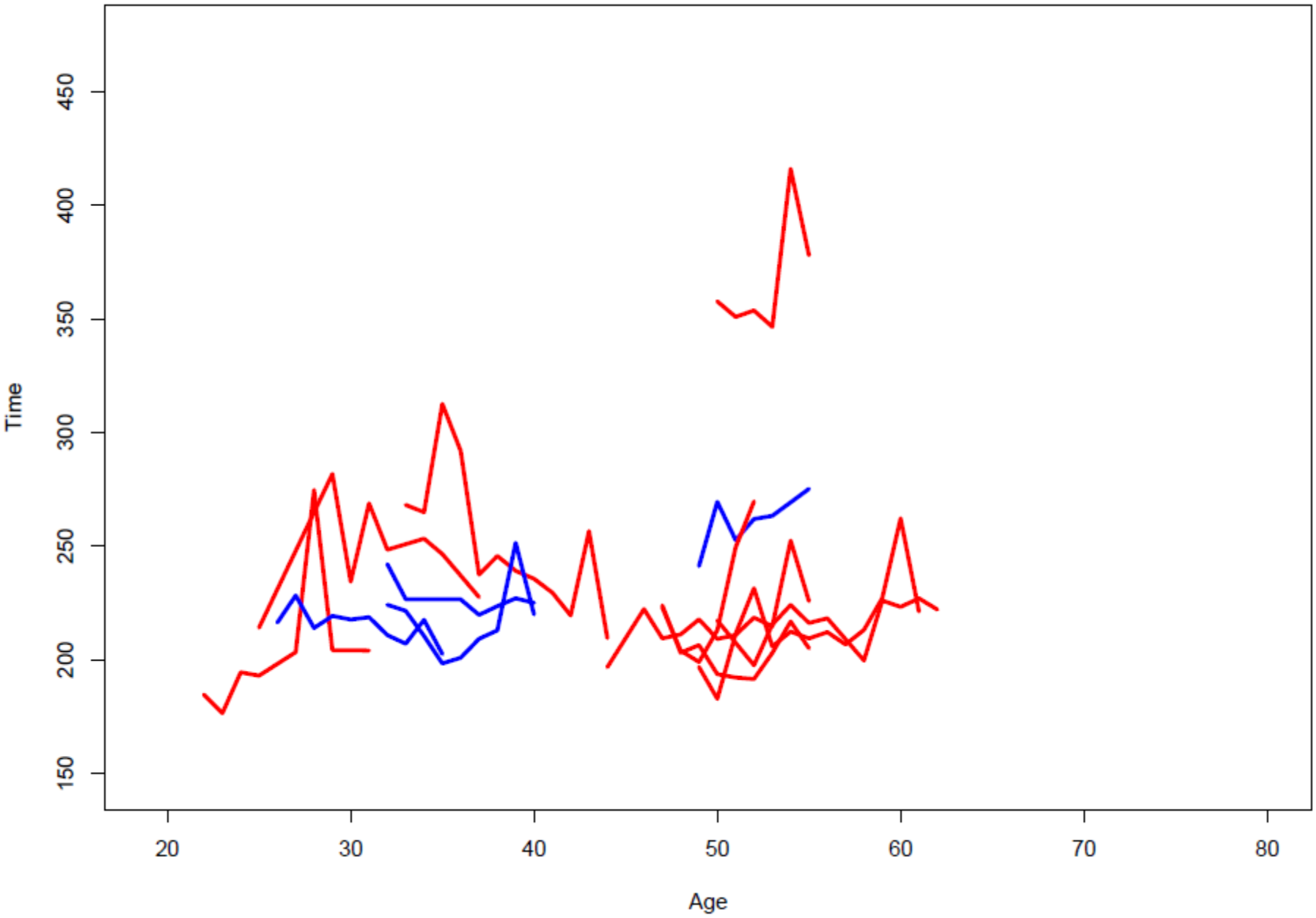

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

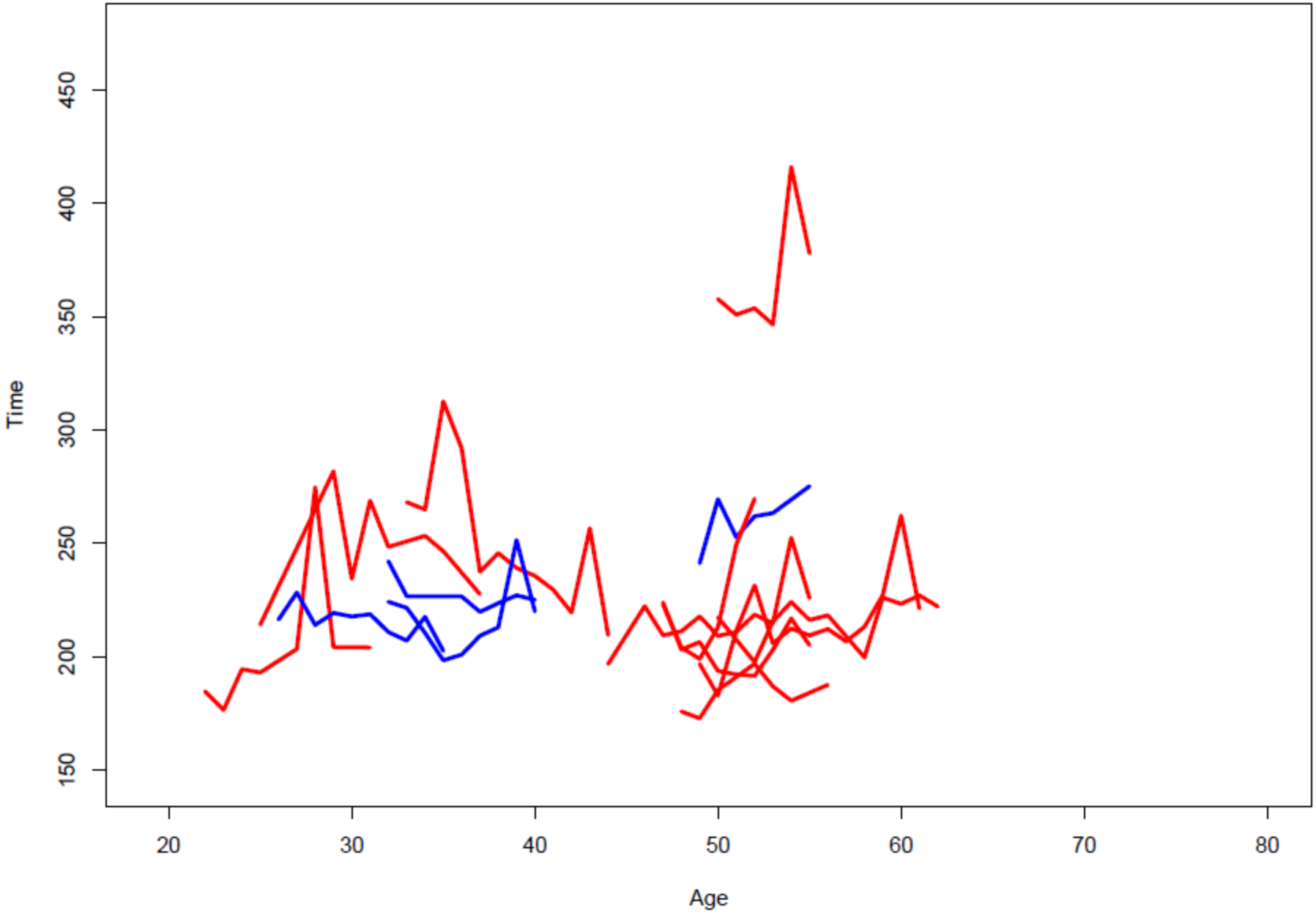

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

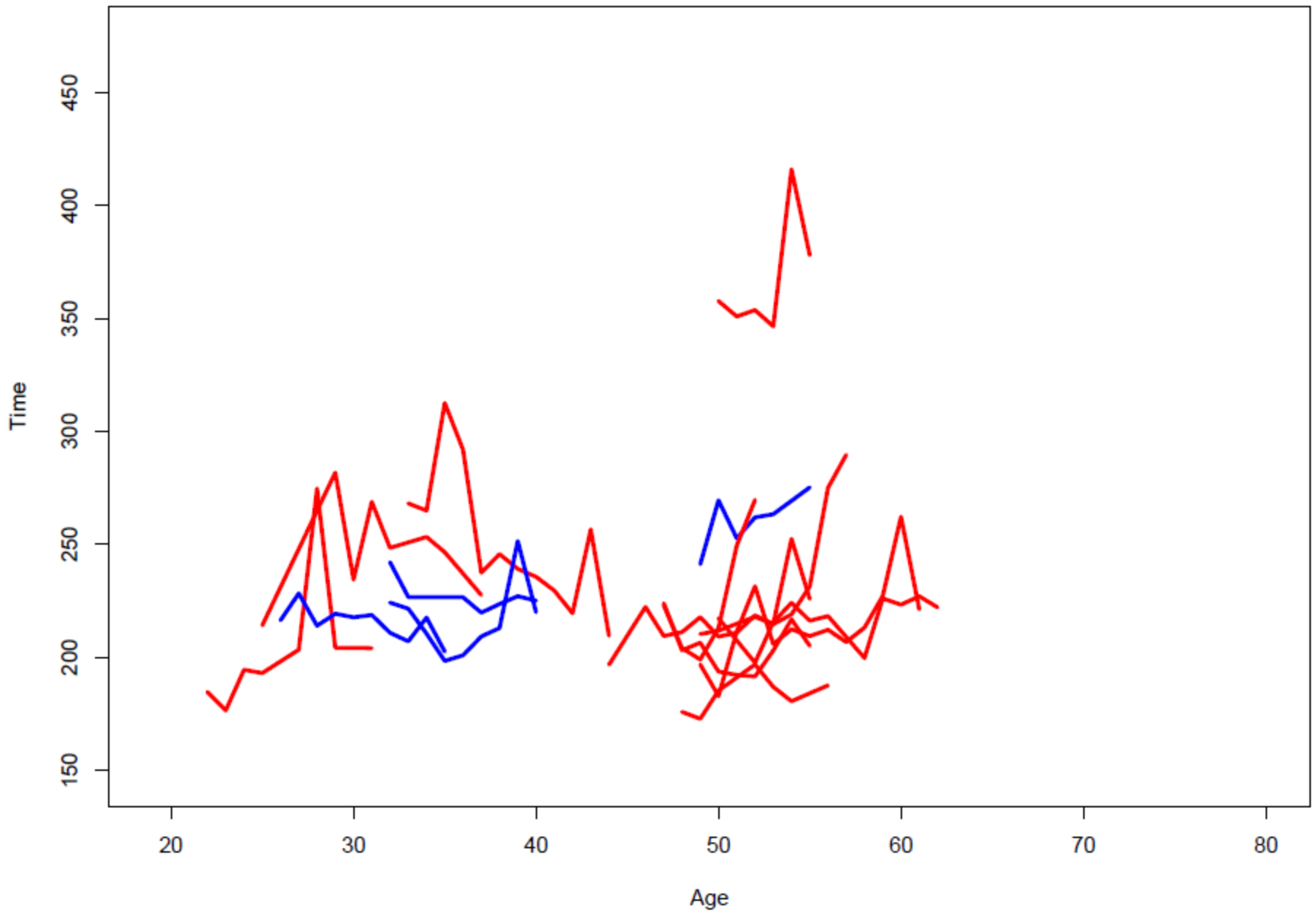

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

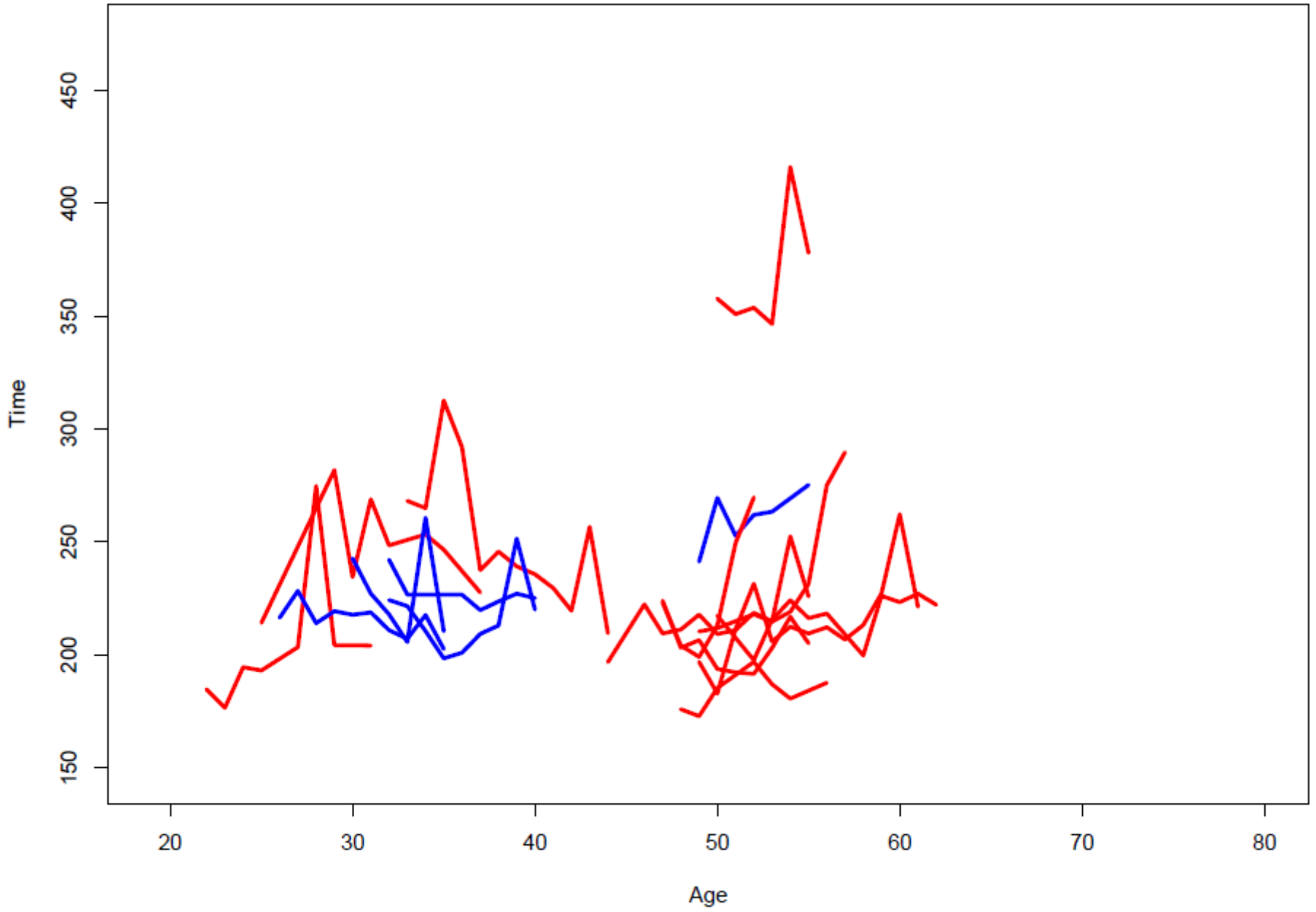

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

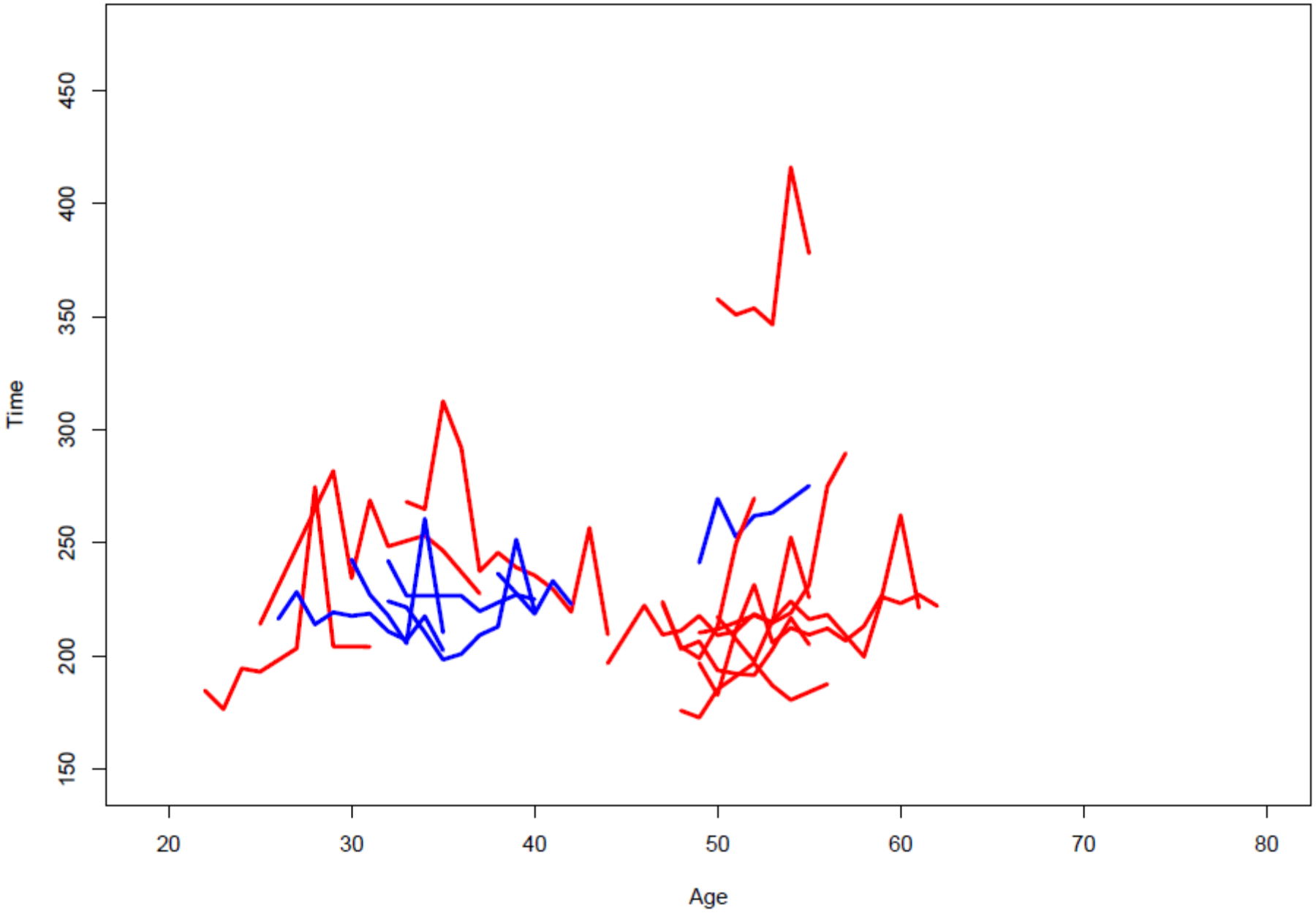

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

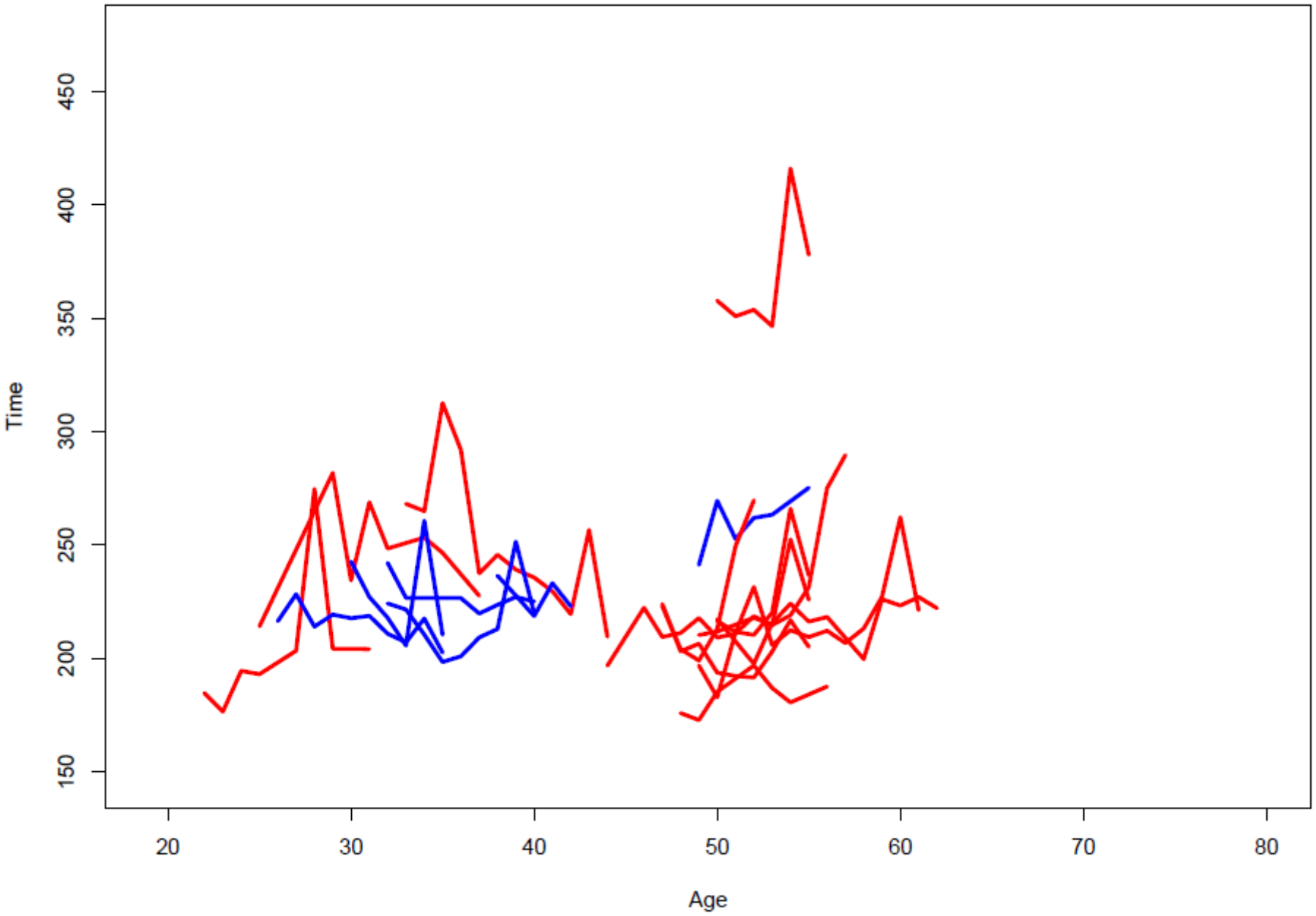

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

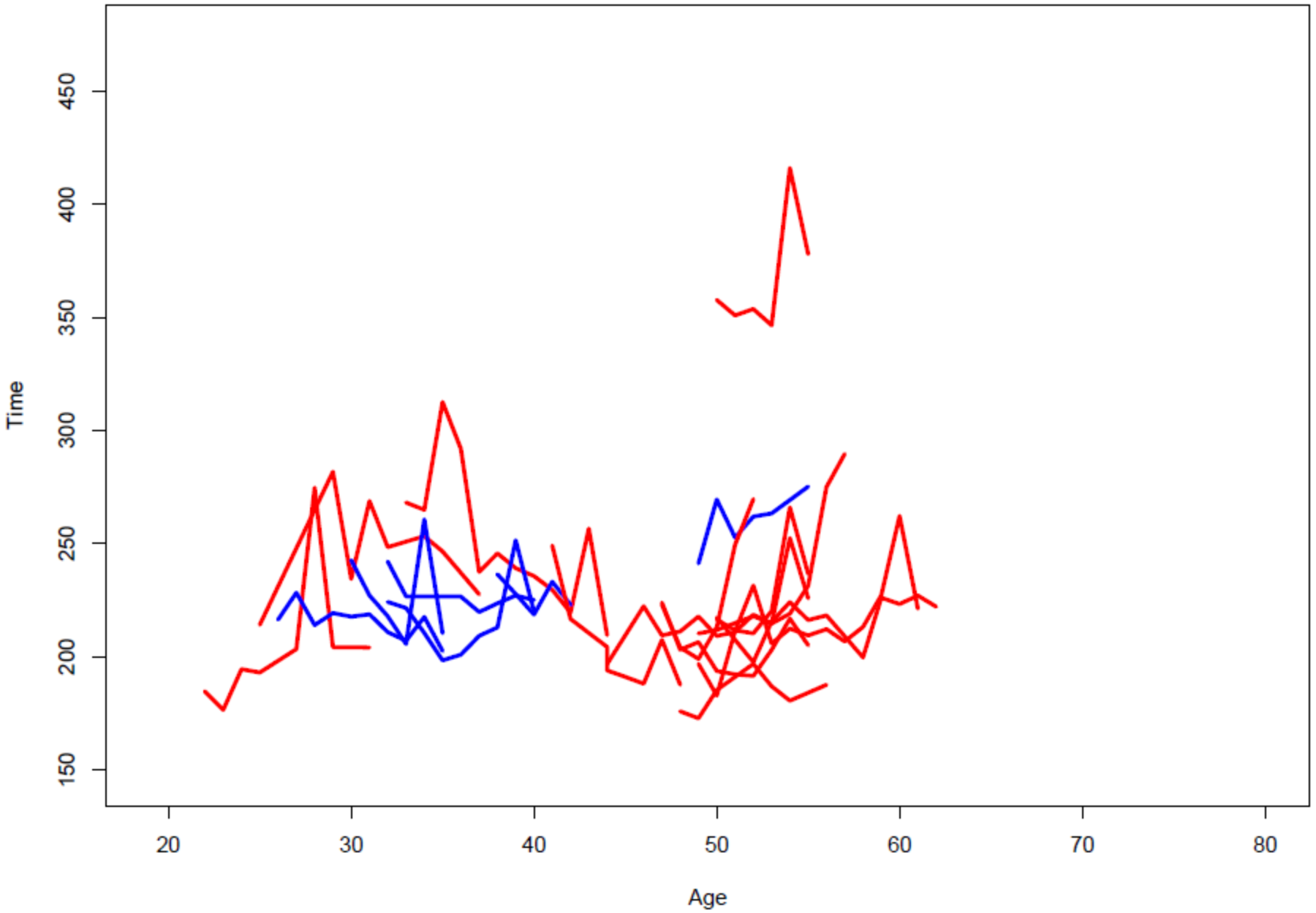

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

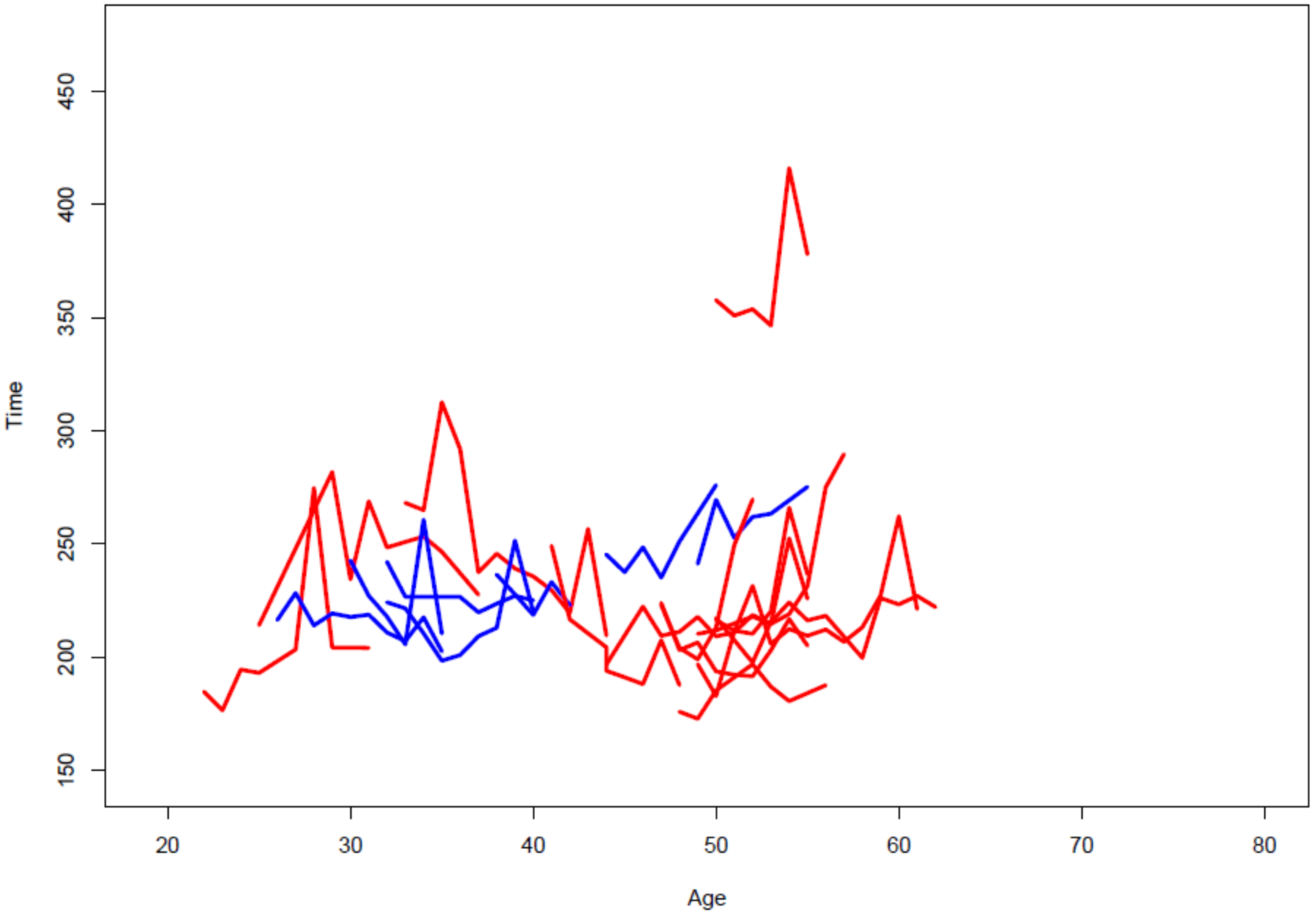

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

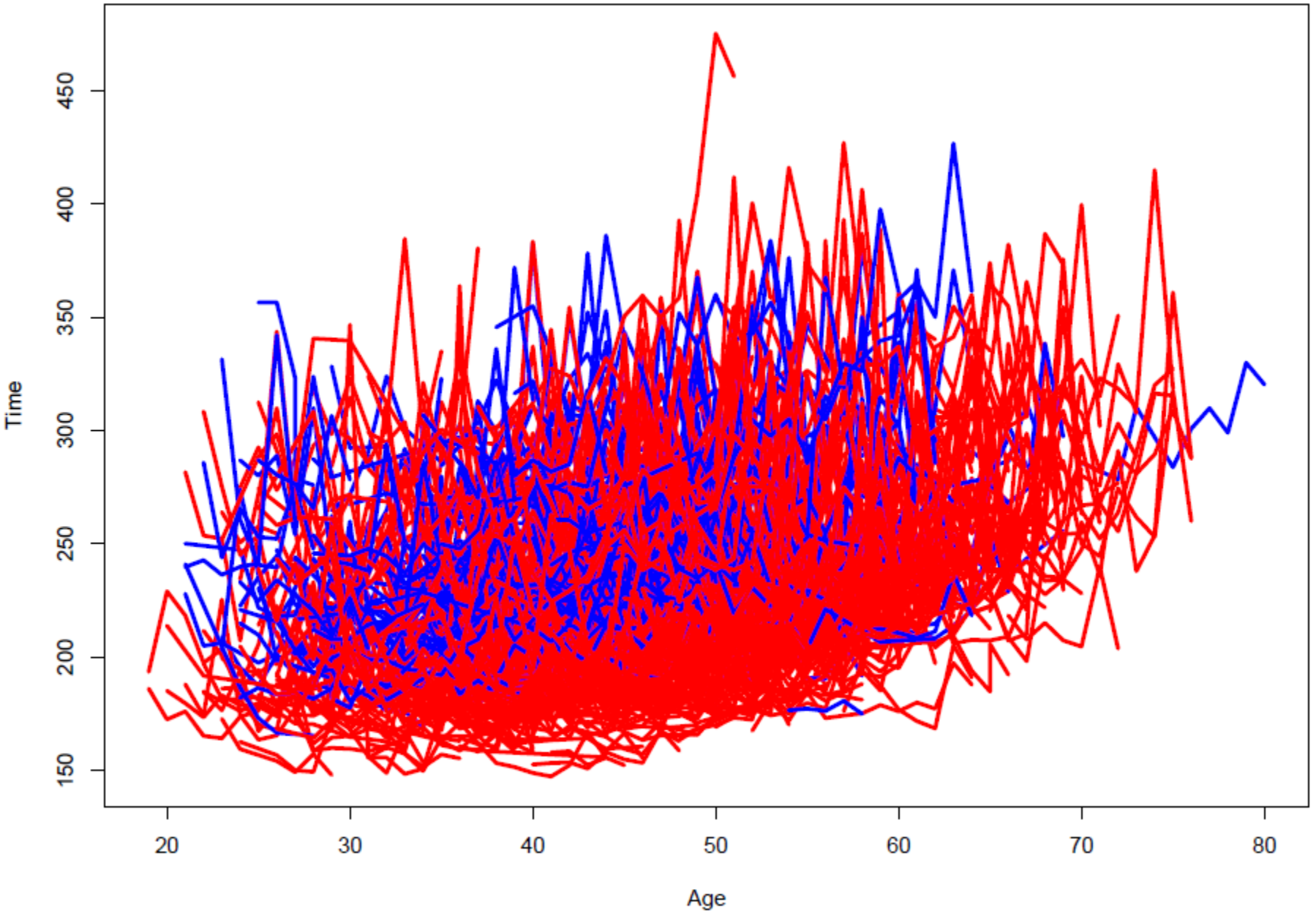

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age


Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age


Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age


Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age


Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age


Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

The Idea

- Each individual runner record is a part-trace of the performance v. age curve for that runner
- Allow for a random “runner effect”
- Also allow for a random “calendar year” effect (2004 and 2012 were very hot)
- Separate men’s and women’s performance
- A refinement (later): also distinguish runners of different ability levels

Statistical Model:

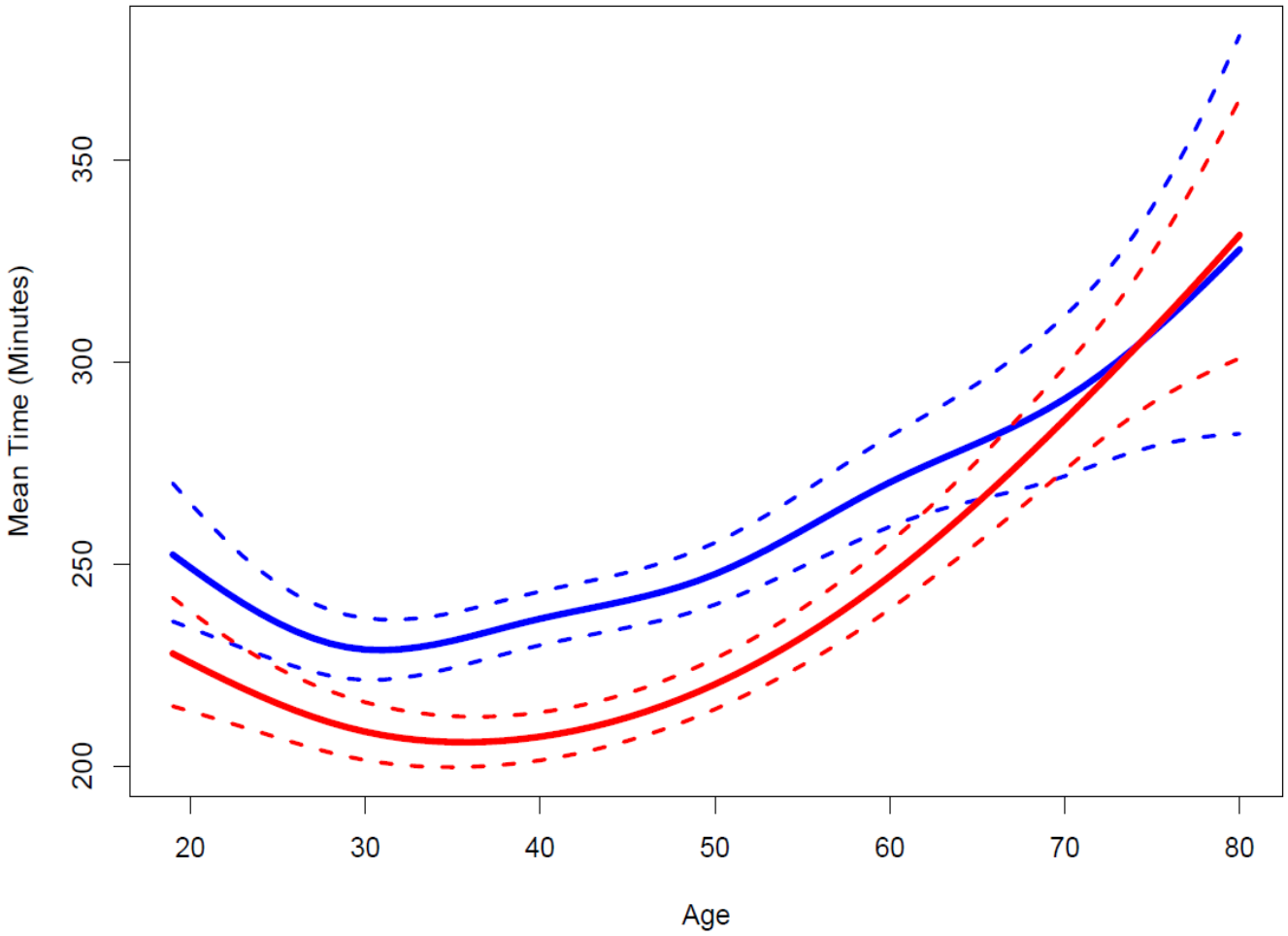
$$\log t_{ij} \;=\; \alpha_i + \beta_{y_{ij}} + S(a_{ij}; K) + \epsilon_{ij},$$

where

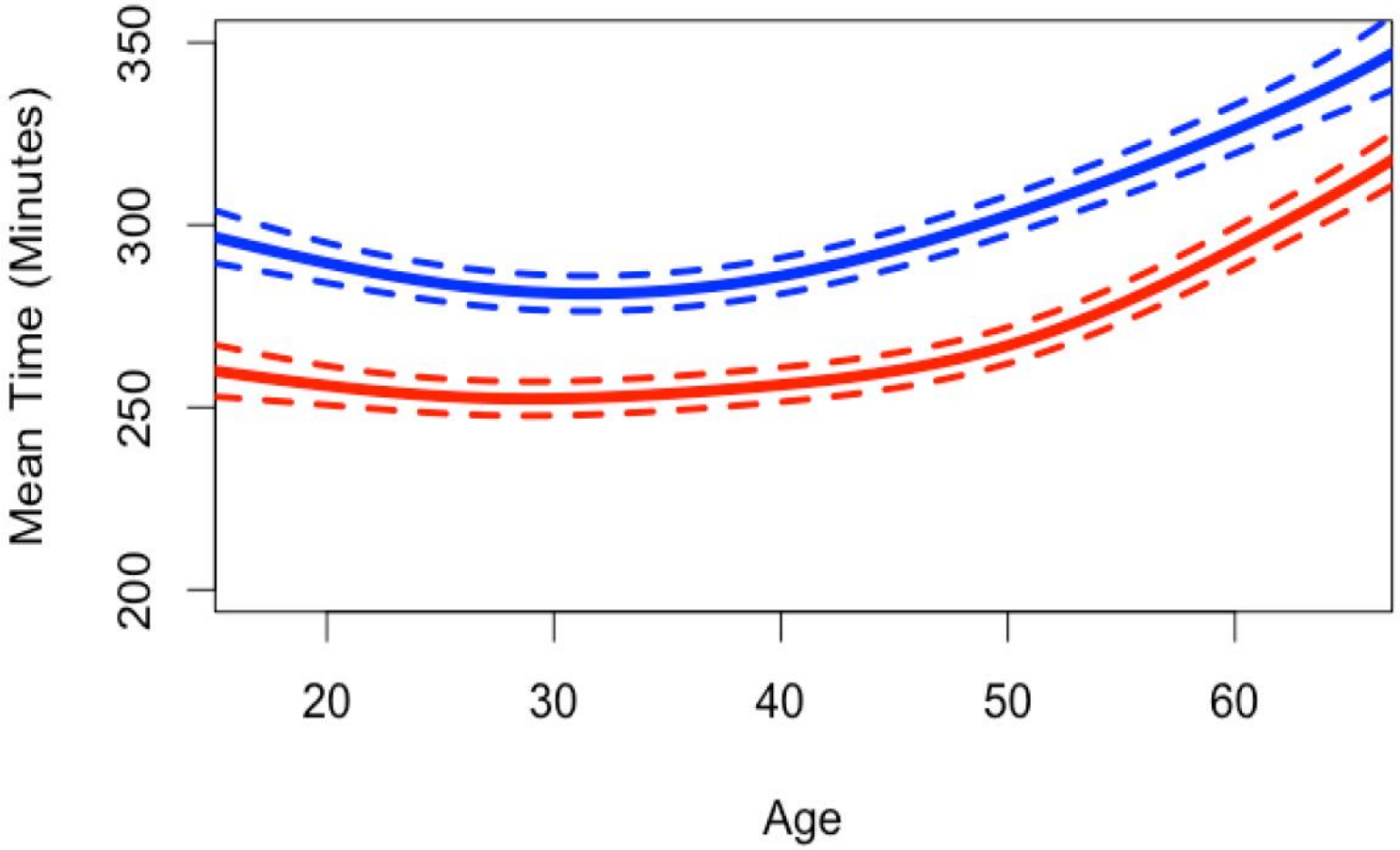
- t_{ij} is the jth finish time of runner i,
- y_{ij} is the year of the jth finish time of runner i,
- a_{ij} is the ith runner's age in her jth finish time,
- α_i represents the overall ability level of runner i (small α_i means a faster runner),
- $\beta_{y_{ij}}$ is a year effect,
- $S(a_{ij}; K)$ represents a nonlinear function of age with K degrees of freedom,
- ϵ_{ij} is a random error.
- Computation: use function `lmer` within R package `lme4`.

Results

- Men's curve (red, with confidence limit)
- Women's curve (blue, with confidence limits)
- Crossover above age 70 almost certainly an artifact
- Other anomalies need to be explained

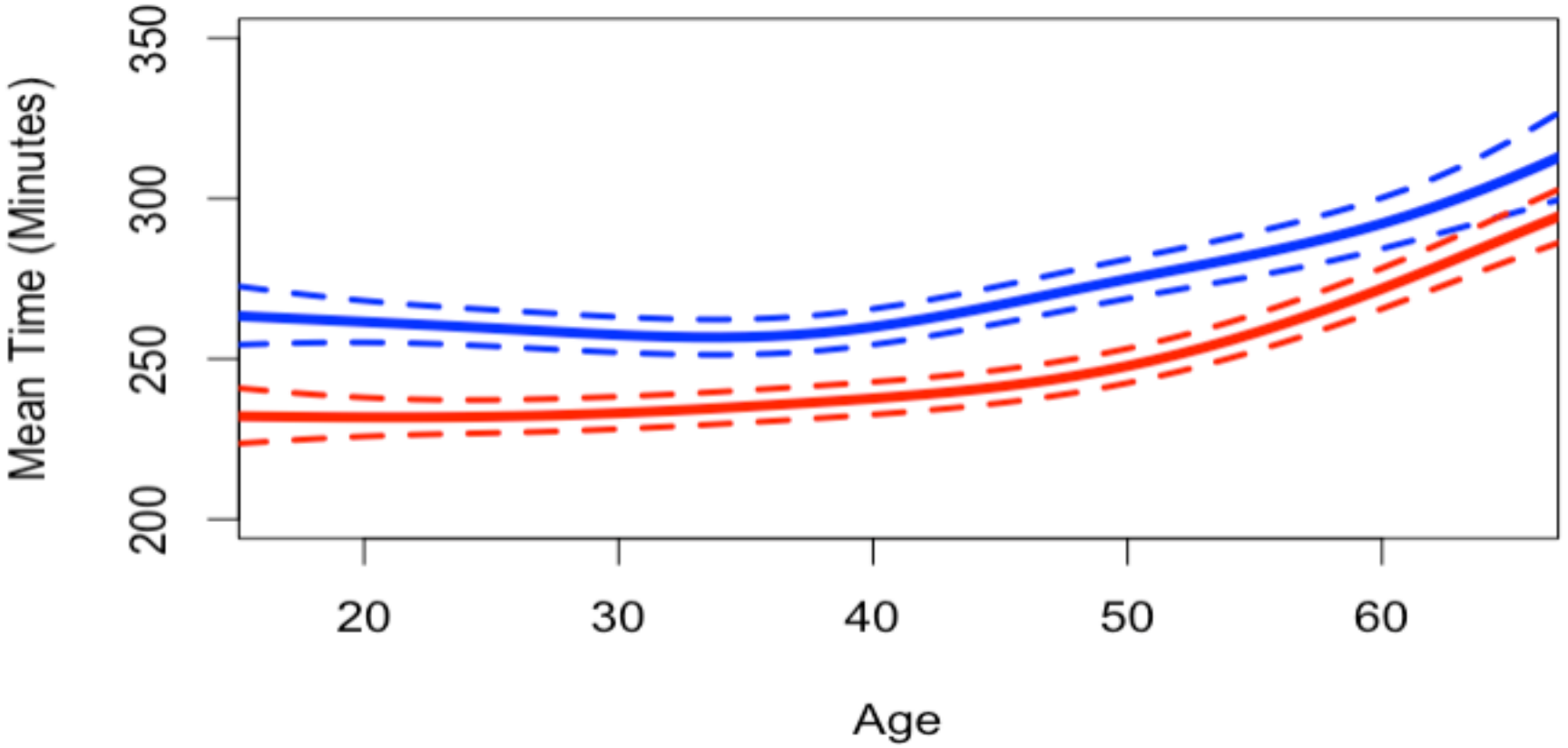


Recent Results (Spencer and Dorit)



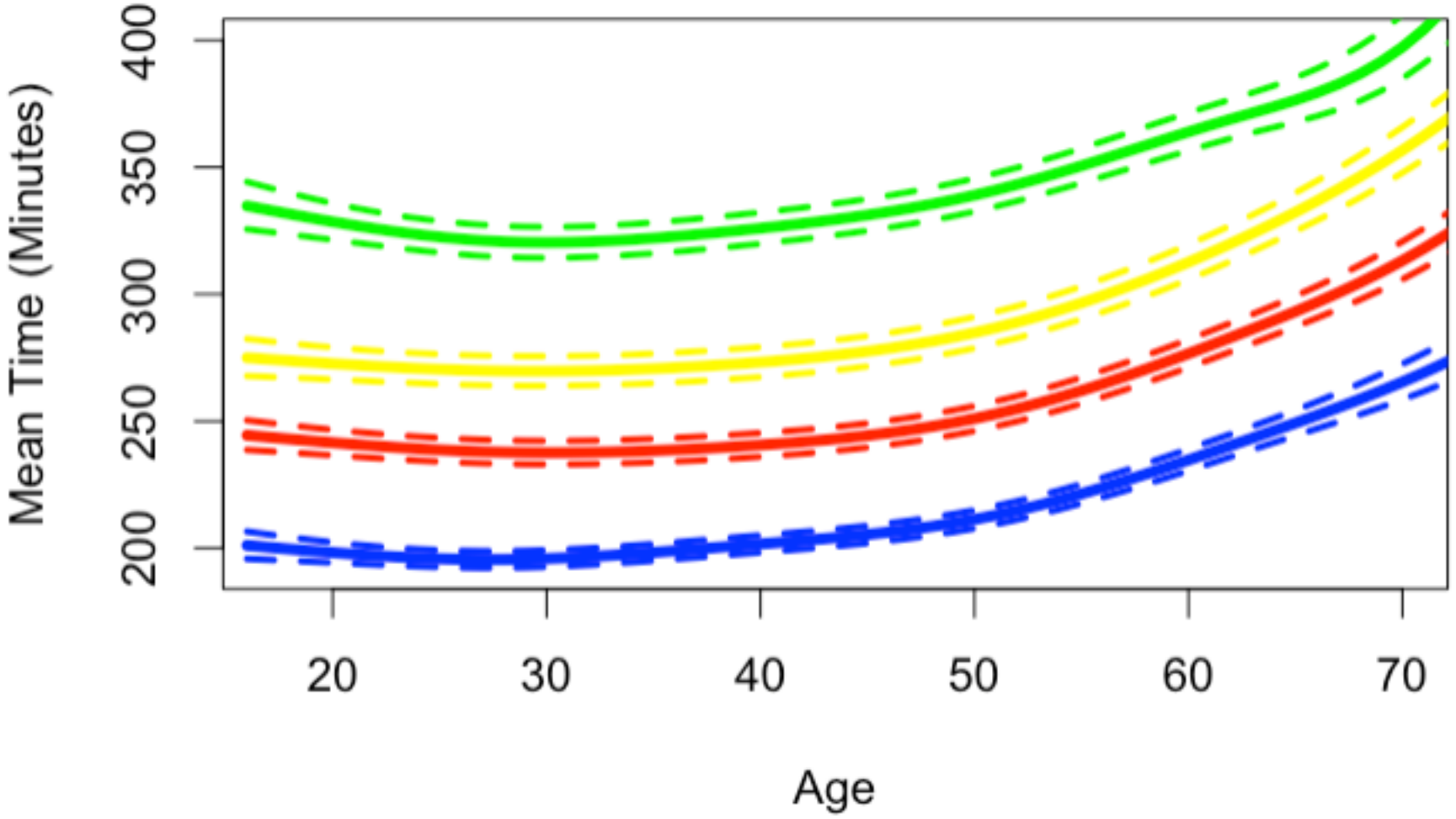
Results from the Chicago Marathon (all runners, 2010-2018), analyzed by the same method

Recent Results (Spencer and Dorit)



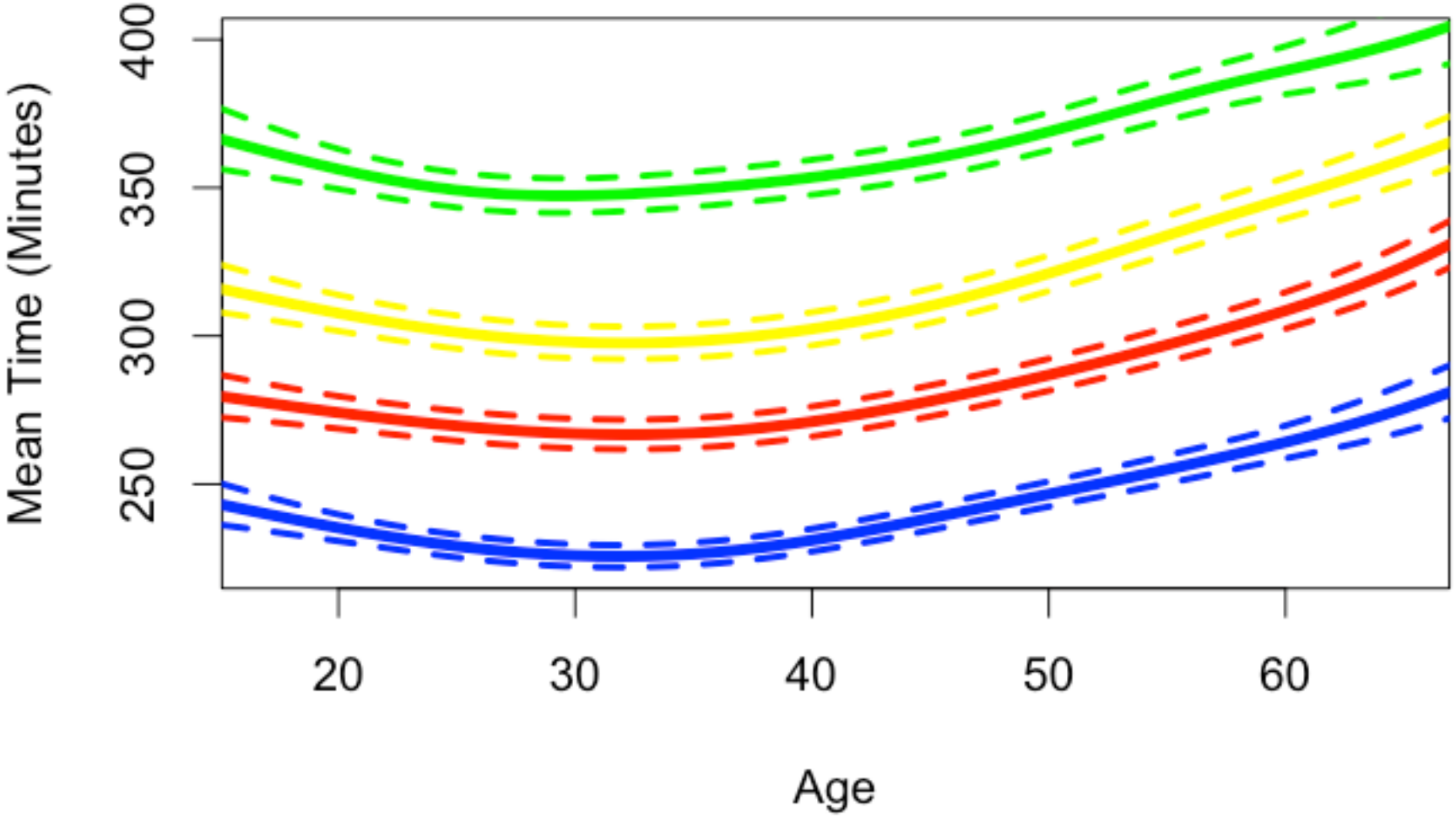
Restricted to runners who have run at least one Boston qualifier time

Recent Results (Spencer and Dorit)

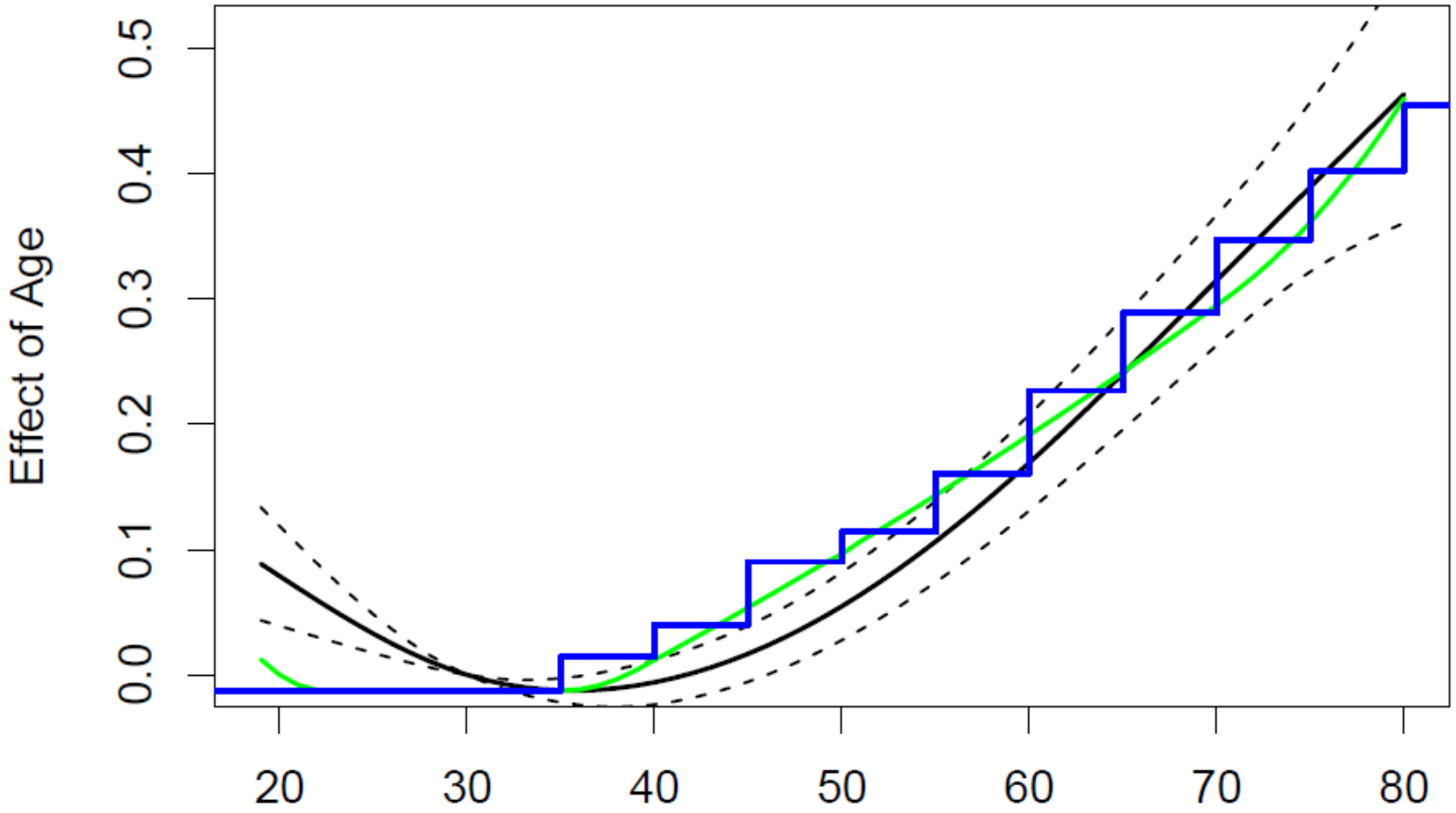


Quartiles for male runners

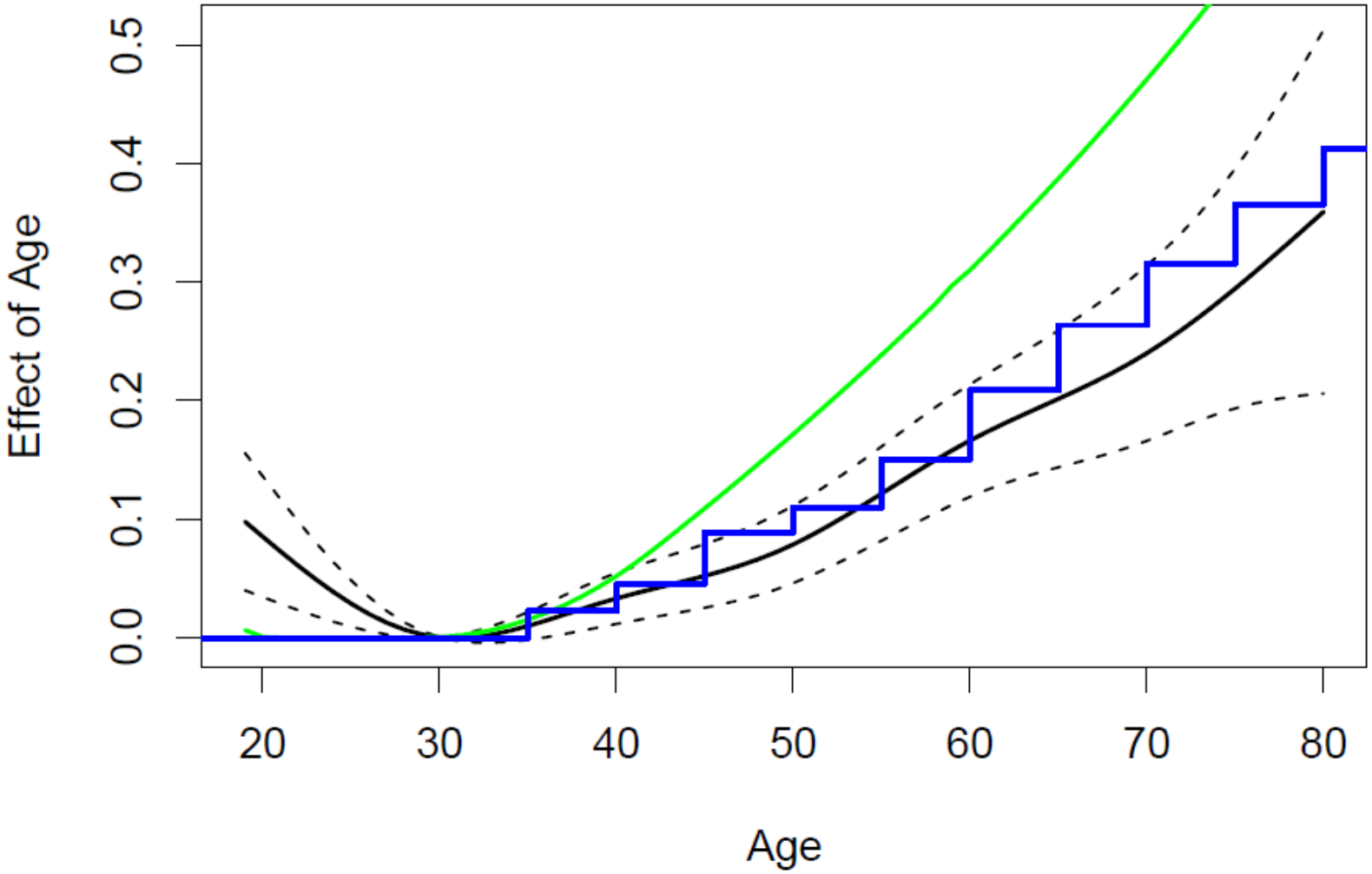
Recent Results (Spencer and Dorit)



Quartiles for female runners

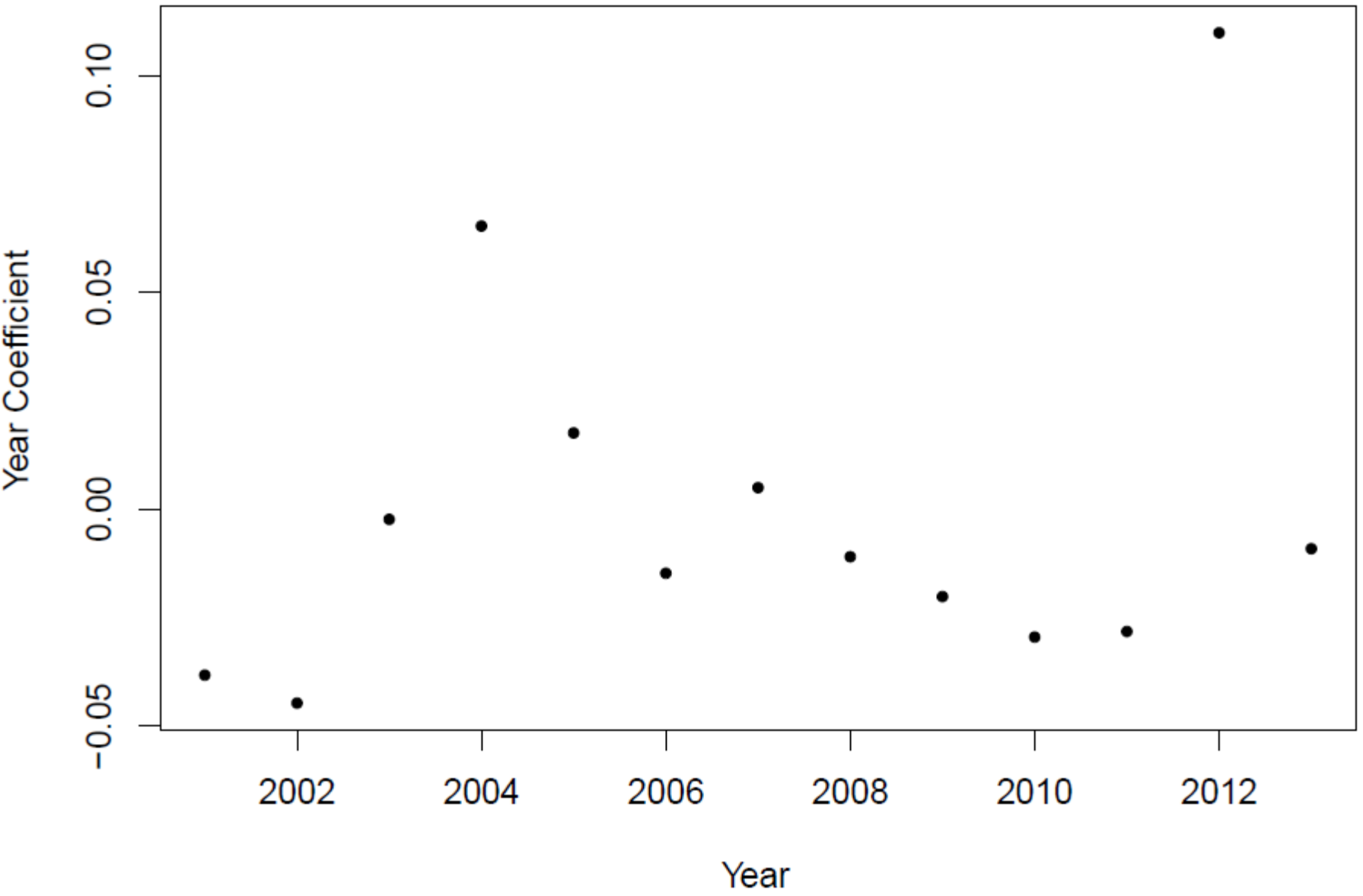


Age-performance curve for men. Solid black curve is spline-based estimate of age effect relative to age 30; dotted black curves are pointwise 95% confidence bounds. Blue step function based on Boston marathon qualifying times; green curve based on age-graded performances. Both the blue step function and the green curve are normalized so that their minimum is the same as that for the estimated spline curve.



Same as previous page, but for women's performances.

Also look at the year effect



We can also consider the implications of these results for Boston Marathon qualifying times

(Warning: These results are quite sensitive to which curve is used and how it is processed)

Age Group	Women			Men		
	Now	Proposed	Age-Graded	Now	Proposed	Age-Graded
35–39	3:40	3:39	3:41	3:10	3:05	3:05
40–44	3:45	3:44	3:52	3:15	3:08	3:12
45–49	3:55	3:49	4:06	3:25	3:13	3:20
50–54	4:00	3:58	4:22	3:30	3:22	3:29
55–59	4:10	4:10	4:41	3:40	3:35	3:39
60–64	4:25	4:22	5:22	3:55	3:53	3:50
65–69	4:40	4:34	5:28	4:10	4:15	4:02
70–74	4:55	4:50	5:57	4:25	4:43	4:16
75–79	5:10	5:14	6:37	4:40	5:15	4:37

Conclusions

- The problem is difficult

Conclusions

- The problem is difficult
- Current Boston Marathon standards have evolved over the years, but no real “science” behind them

Conclusions

- The problem is difficult
- Current Boston Marathon standards have evolved over the years, but no real “science” behind them
- Age-graded standards are popular in the running community, but they have problems

Conclusions

- The problem is difficult
- Current Boston Marathon standards have evolved over the years, but no real “science” behind them
- Age-graded standards are popular in the running community, but they have problems
- The longitudinal method proposed here is the only method based on how individual runners’ performances change with age

Conclusions

- The problem is difficult
- Current Boston Marathon standards have evolved over the years, but no real “science” behind them
- Age-graded standards are popular in the running community, but they have problems
- The longitudinal method proposed here is the only method based on how individual runners’ performances change with age
- Intention for future work: we plan to look at other marathons and alternative statistical methods

Conclusions

- The problem is difficult
- Current Boston Marathon standards have evolved over the years, but no real “science” behind them
- Age-graded standards are popular in the running community, but they have problems
- The longitudinal method proposed here is the only method based on how individual runners’ performances change with age
- Intention for future work: we plan to look at other marathons and alternative statistical methods
- Thank you for your attention!