# *THE VARIATION OF MARATHON PERFORMANCES WITH AGE AND SEX*

**Richard L Smith**

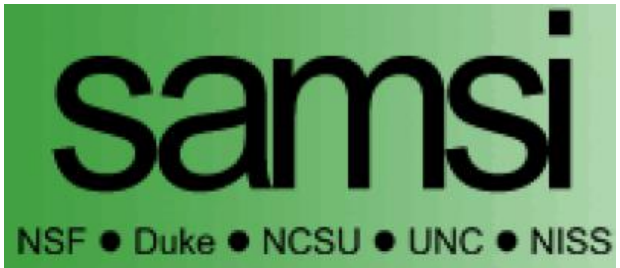
**University of North Carolina and SAMSI**

with input from:

Scott Powers (Stanford)

Jessi Cisewski (Carnegie Mellon)

Jack Fleming and Michael Pieroni (BAA)



UNC
COLLEGE OF
ARTS & SCIENCES

***PART I:*** Our work on qualifying times for the Boston Marathon (Smith, Powers, Cisewski, *Chance*, to appear)

***PART II:*** A new "longitudinal" approach to age-related performances (work in progress)

# *BACKGROUND*

- The Boston Marathon is the only major marathon to require qualifying standards of (almost) all participants
- The need to revise the standards became apparent after entries for the 2011 race closed out in one day
- My interest in this arose out of extensive press discussion, including the idea that women's standards should be tightened relative to the men's
- I sent the BAA a proposal to do some statistical analysis
- Much to my surprise, they accepted …

# 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 |

# *Our Work with the BAA*

- Objective: try to determine the total number of eligible runners under existing standard and under various proposed modifications
- Total of 641 “feeder” races from which at least one person qualified for the 2010 Boston Marathon
- Eliminated 362 as being small races, or out of US/Canada, or out of qualification window
- Left 279 races from which we would estimate qualifiers
- Complete counts from 37 “top feeder” races that sent most qualifiers to the Boston Marathon
- Complete counts from 5 other large races
- Randomly sampled 25 races out of remaining 237, scaled up numbers to estimate total qualifiers in those 237 races

**Table 2: Estimated totals by age group of all male and female marathon runners in all races, and numbers of qualifiers under 2011 qualifying standards.**

| Age Group | Total Runners | | Boston Qualifiers | |
|---|---|---|---|---|
| | M | F | M | F |
| 18-34 | 126780 | 126861 | 12004 | 12782 |
| 35-39 | 67271 | 48298 | 6630 | 5787 |
| 40-44 | 70890 | 45339 | 8305 | 6594 |
| 45-49 | 60644 | 34101 | 10994 | 6783 |
| 50-54 | 45712 | 22033 | 7541 | 3420 |
| 55-59 | 27609 | 10453 | 4378 | 1533 |
| 60-64 | 15799 | 4686 | 3204 | 649 |
| 65-69 | 6270 | 1584 | 1167 | 192 |
| 70-74 | 2488 | 397 | 560 | 52 |
| 75-79 | 601 | 113 | 84 | 9 |
| 80-99 | 148 | 28 | 21 | 6 |
| All ages | 424210 | 293893 | 54888 | 37809 |
| Total | | 718106 | | 92695 |

**Table 3: Percentage of qualifiers for all races with QT adjustments**

| Age | QT1 | | QT1 - 2 | | QT1 - 5 | | QT1 - 10 | | QT1 - 15 | | QT1 - 20 | | QT1 – 30 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Group | M | F | M | F | M | F | M | F | M | F | M | F | M | F |
| 18-34 | 9.5 | 10.1 | 8.2 | 8.9 | 6.7 | 7.0 | 5.3 | 5.1 | 3.6 | 3.5 | 2.7 | 2.5 | 1.4 | 1.2 |
| 35-39 | 9.9 | 12.0 | 8.4 | 10.6 | 6.8 | 8.4 | 4.8 | 6.0 | 3.6 | 4.3 | 2.2 | 3.0 | 0.9 | 1.5 |
| 40-44 | 11.7 | 14.5 | 10.0 | 13.0 | 8.2 | 10.7 | 5.9 | 7.7 | 4.3 | 5.4 | 3.1 | 3.8 | 1.0 | 1.7 |
| 45-49 | 18.1 | 19.9 | 15.6 | 17.9 | 12.6 | 14.9 | 9.3 | 11.6 | 6.8 | 8.2 | 5.0 | 6.1 | 2.4 | 3.0 |
| 50-54 | 16.5 | 15.5 | 14.6 | 14.1 | 11.9 | 12.1 | 8.4 | 8.7 | 5.9 | 6.3 | 3.9 | 4.3 | 1.7 | 1.8 |
| 55-59 | 15.9 | 14.7 | 14.1 | 13.3 | 11.8 | 11.3 | 8.6 | 9.0 | 6.0 | 6.3 | 4.1 | 4.2 | 1.7 | 2.2 |
| 60-64 | 20.3 | 13.8 | 18.1 | 12.4 | 15.3 | 11.2 | 11.5 | 9.0 | 8.8 | 7.5 | 6.7 | 5.9 | 3.5 | 3.7 |
| 65-69 | 18.6 | 12.1 | 17.0 | 11.7 | 14.6 | 9.8 | 12.1 | 8.5 | 10.0 | 6.7 | 7.9 | 5.9 | 4.5 | 3.4 |
| 70-74 | 22.5 | 13.1 | 20.8 | 12.6 | 17.4 | 12.1 | 13.6 | 9.8 | 11.5 | 7.6 | 8.8 | 6.8 | 5.2 | 5.0 |
| 75-79 | 14.0 | 8.0 | 13.6 | 8.0 | 12.8 | 7.1 | 11.5 | 6.2 | 9.8 | 4.4 | 8.3 | 3.5 | 6.2 | 2.7 |
| 80-99 | 14.2 | 21.4 | 13.5 | 21.4 | 12.2 | 21.4 | 12.2 | 21.4 | 11.5 | 21.4 | 8.1 | 21.4 | 7.4 | 21.4 |

# *Conclusions from this exercise*

- In 18-34 age group, current standards *very slightly* favor women, but this would be reversed with as little as a 2-minute adjustment to women's qualifying times
- Qualifying percentages are higher in older age groups, where there are far fewer competitors (*is this fair?* – question for second part of talk)
- In my opinion, women over 60 do get a raw deal – there are far fewer of them than men in the same age groups, but the qualifying percentages are substantially smaller than men's

# *Decision by the BAA*

- Reduced all qualifying times by 5 minutes but also eliminated 59-second cushion
- Introduced 4-tier entry system which accepted runners in order of margin under the qualifying time for their age/sex group
- 2012: Old qualifying times still in effect, accepted all runners 1 m. 14 s. or better under their qualifying time
- 2013: New qualifying times in effect, accepted all qualified runners
- 2014: Expanded field included non-finishers from 2013; accepted all runners 1 m. 38 s. or better under their qualifying time (mine was 1 m. 54 s. under QT)

# *How well did our projections hold up?*

# Figure 1: Numbers of estimated and actual qualifiers by time increment

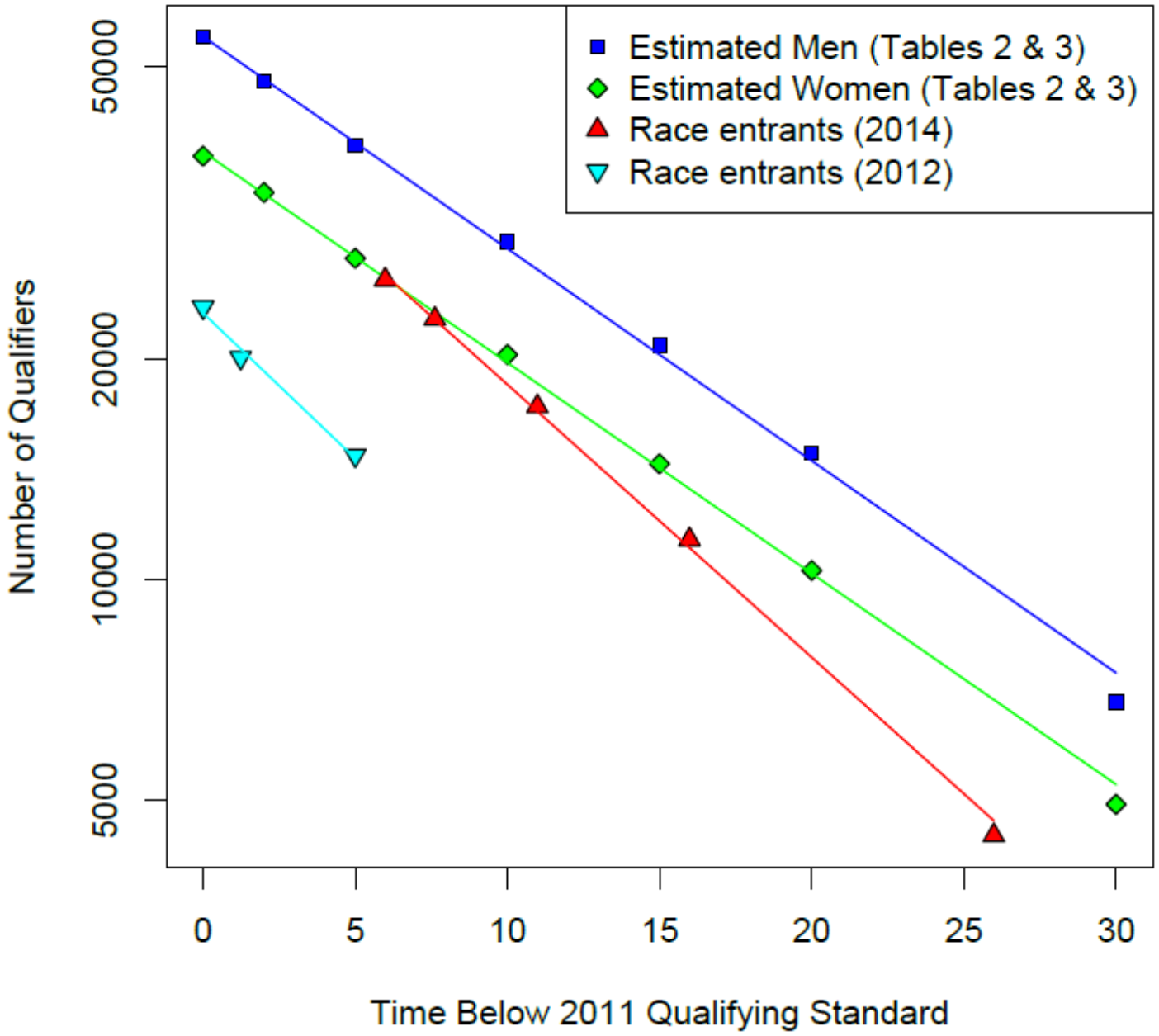


**Table 9: Changes in Expected and Observed Percentages from Old to New Qualifying Times for Each Age-Sex Category**

| Age | Expected | | Observed | |
|---|---|---|---|---|
| | Men | Women | Men | Women |
| 18-34 | -0.093 | -0.551 | -0.183 | -0.193 |
| 35-39 | -0.305 | -0.113 | -0.487 | 0.528 |
| 40-44 | -0.310 | 0.087 | -0.725 | -0.596 |
| 45-49 | -0.318 | 0.392 | -0.868 | -0.286 |
| 50-54 | 0.040 | 0.324 | -0.032 | 0.357 |
| 55-59 | 0.110 | 0.156 | 0.588 | 0.444 |
| 60-64 | 0.168 | 0.110 | 0.380 | 0.249 |
| 65-69 | 0.146 | 0.036 | 0.495 | 0.120 |
| 70-74 | 0.057 | 0.021 | 0.076 | 0.011 |
| 75-79 | 0.031 | 0.003 | 0.075 | 0.002 |
| 80 and over | 0.006 | 0.003 | 0.034 | 0.011 |

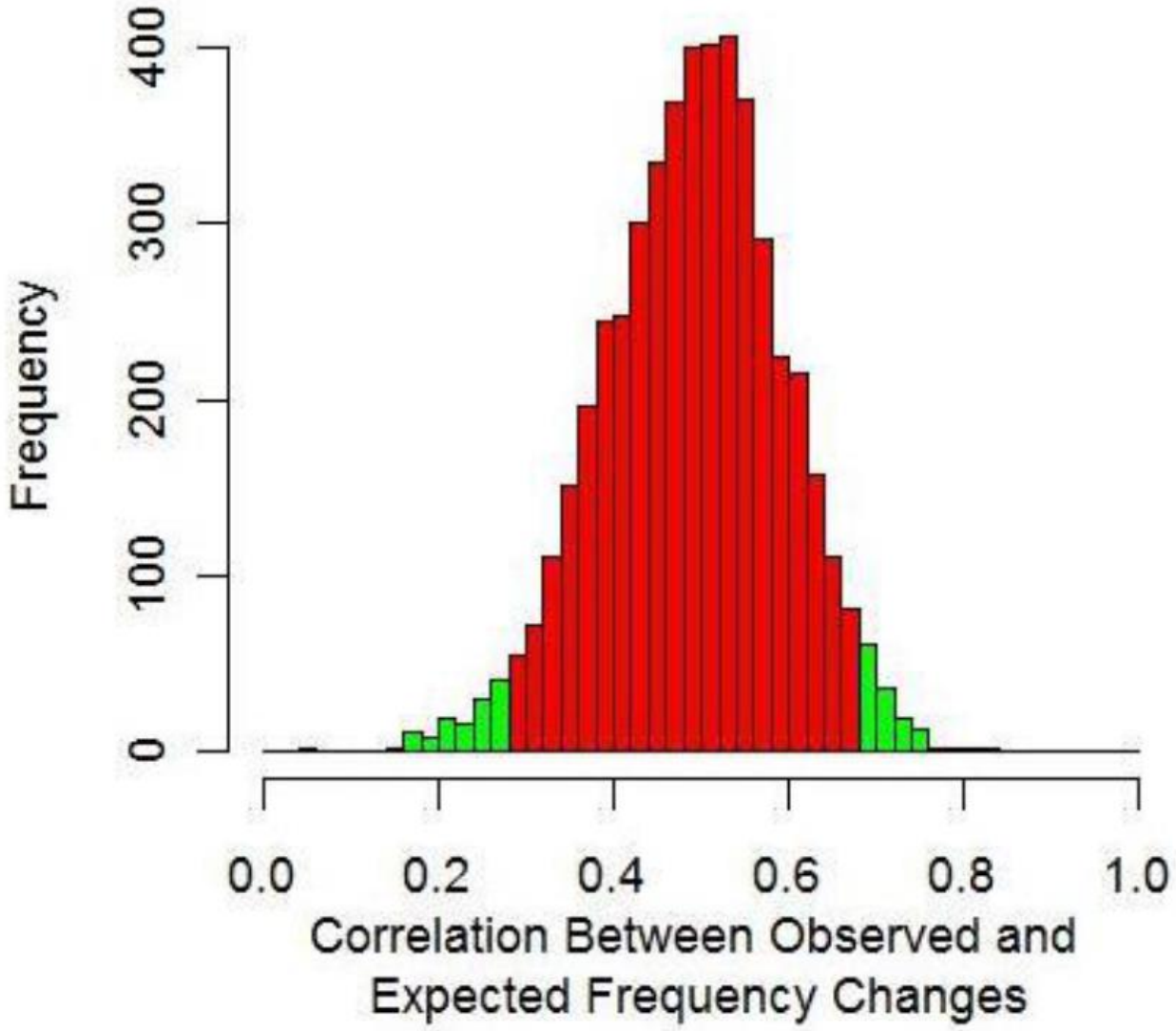
***Correlation 0.527***

# *Is This Correlation Statistically Significant?*

# *Is This Correlation Statistically Significant? Run a bootstrap …*

# *Is This Correlation Statistically Significant? Run a bootstrap …*

**Figure 2: Histogram of Bootstrapped Correlation Coefficients: Central 95% in Red**



# *Part II: Longitudinal Approach to the Performance v. Age Problem*

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

# *Part II: 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

# *Part II: 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

# *Part II: Longitudinal Approach to the Performance v. Age Problem*

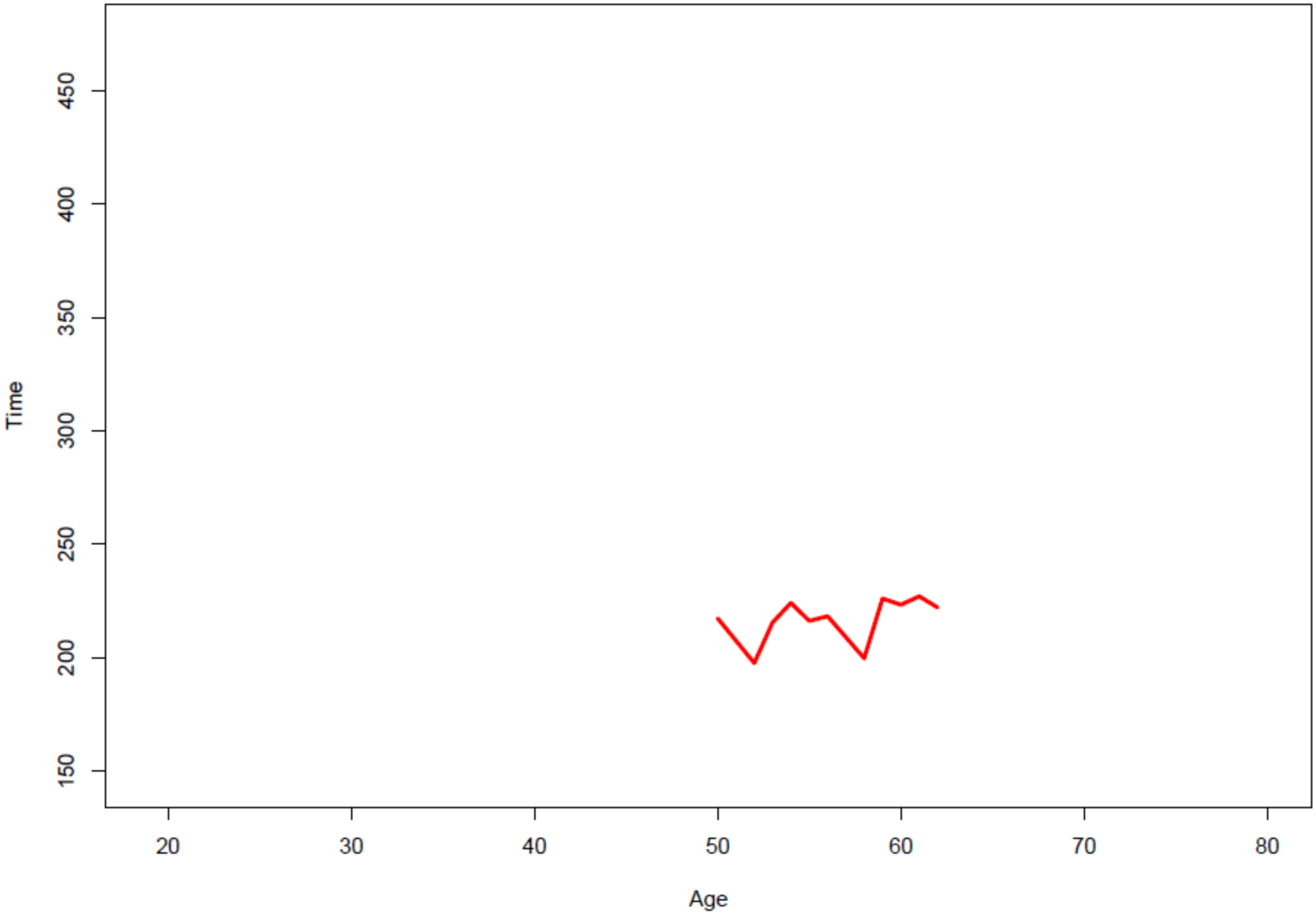
- About 500 runners have run the Boston marathon at least 10 years in succession (BAA)
- No easy way to reconstruct the list
- I used the datasets I had available to find 1,272 runners who had run each of 2010, 2011, 2013
- I then used the BAA archive to find all of those runners who had run at least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

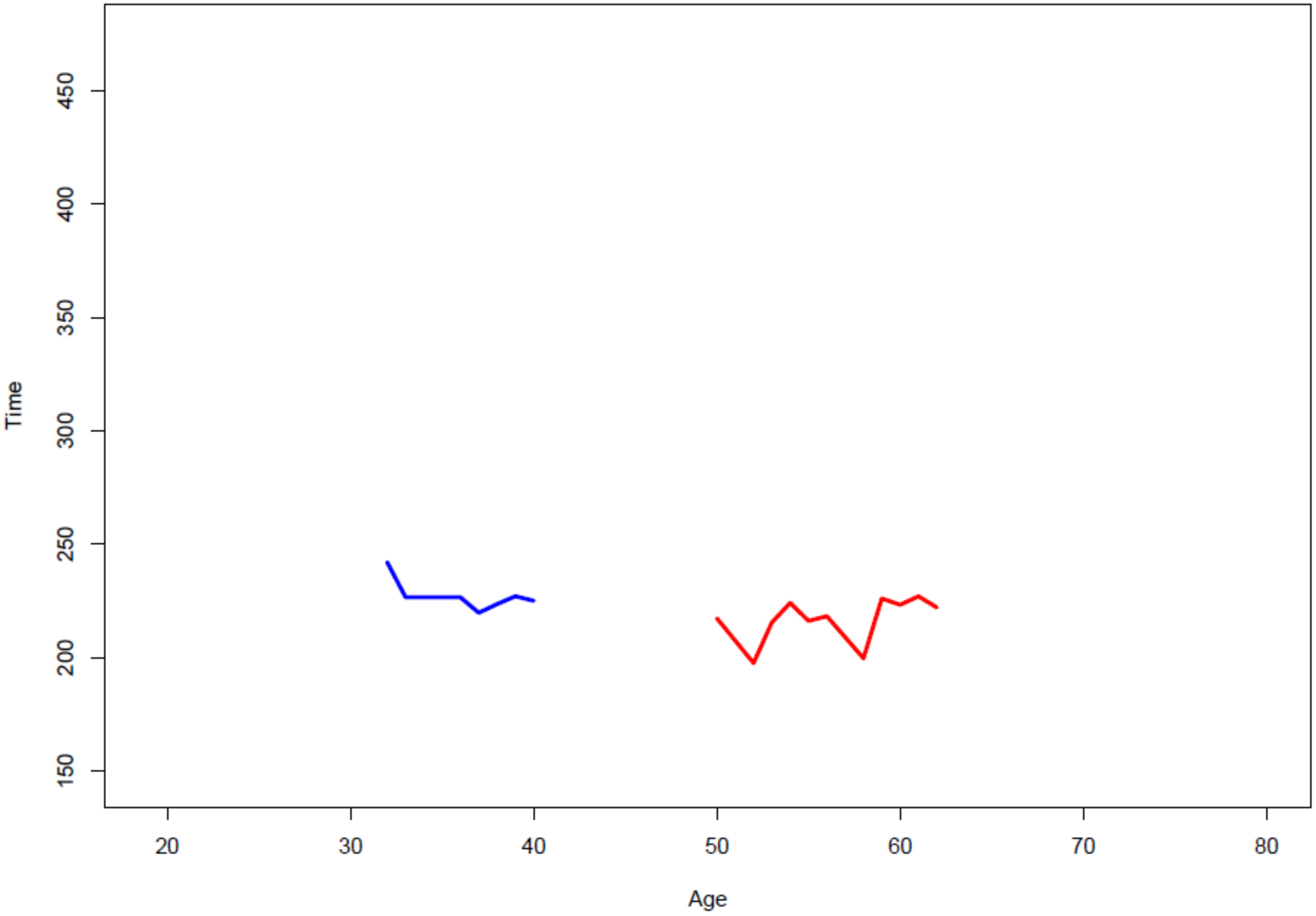
# *Part II: Longitudinal Approach to the Performance v. Age Problem*

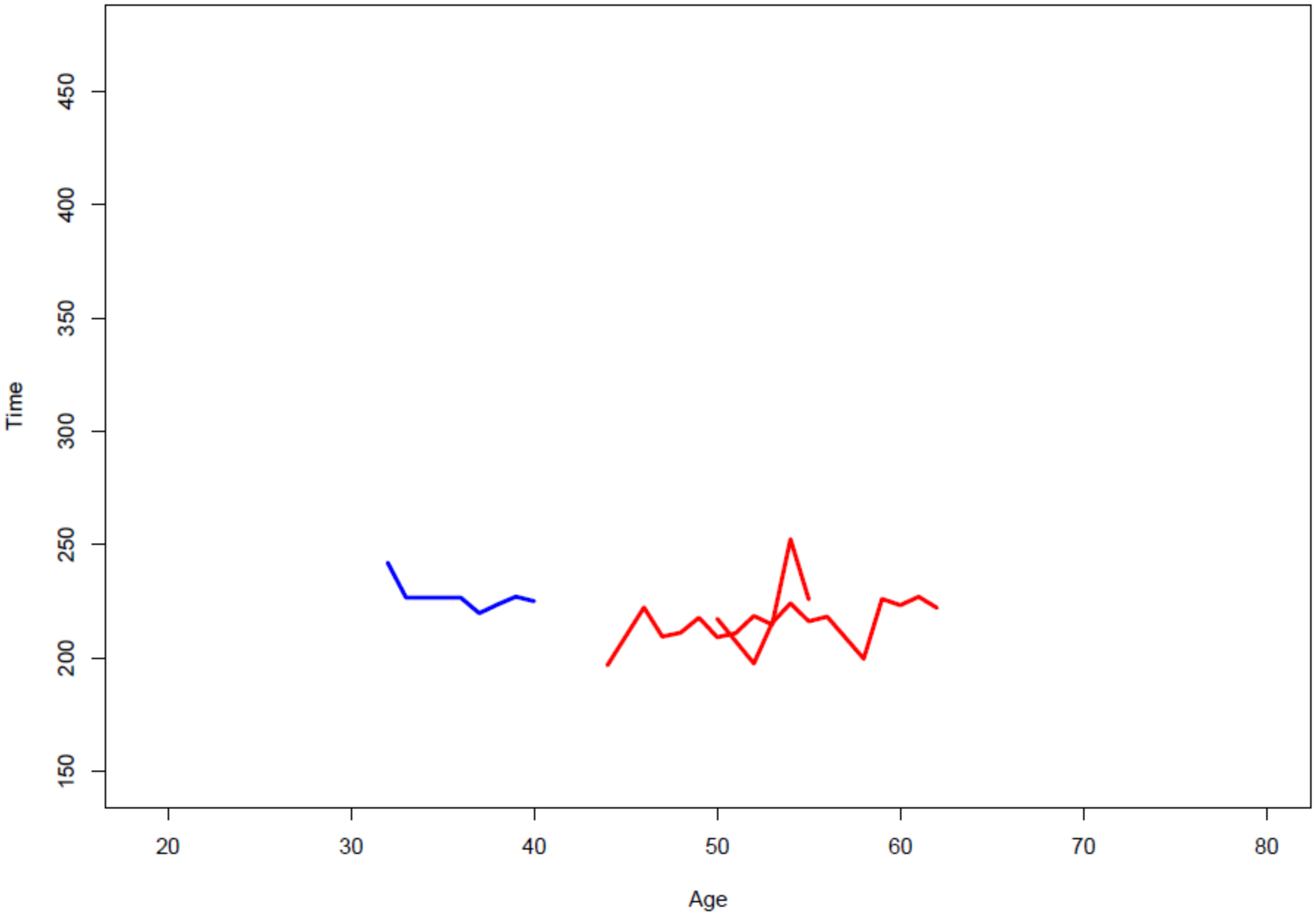
- About 500 runners have run the Boston marathon at least 10 years in succession (BAA)
- No easy way to reconstruct the list
- I used the datasets I had available to find 1,272 runners who had run each of 2010, 2011, 2013
- I then used the BAA archive to find all of those runners who had run at least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

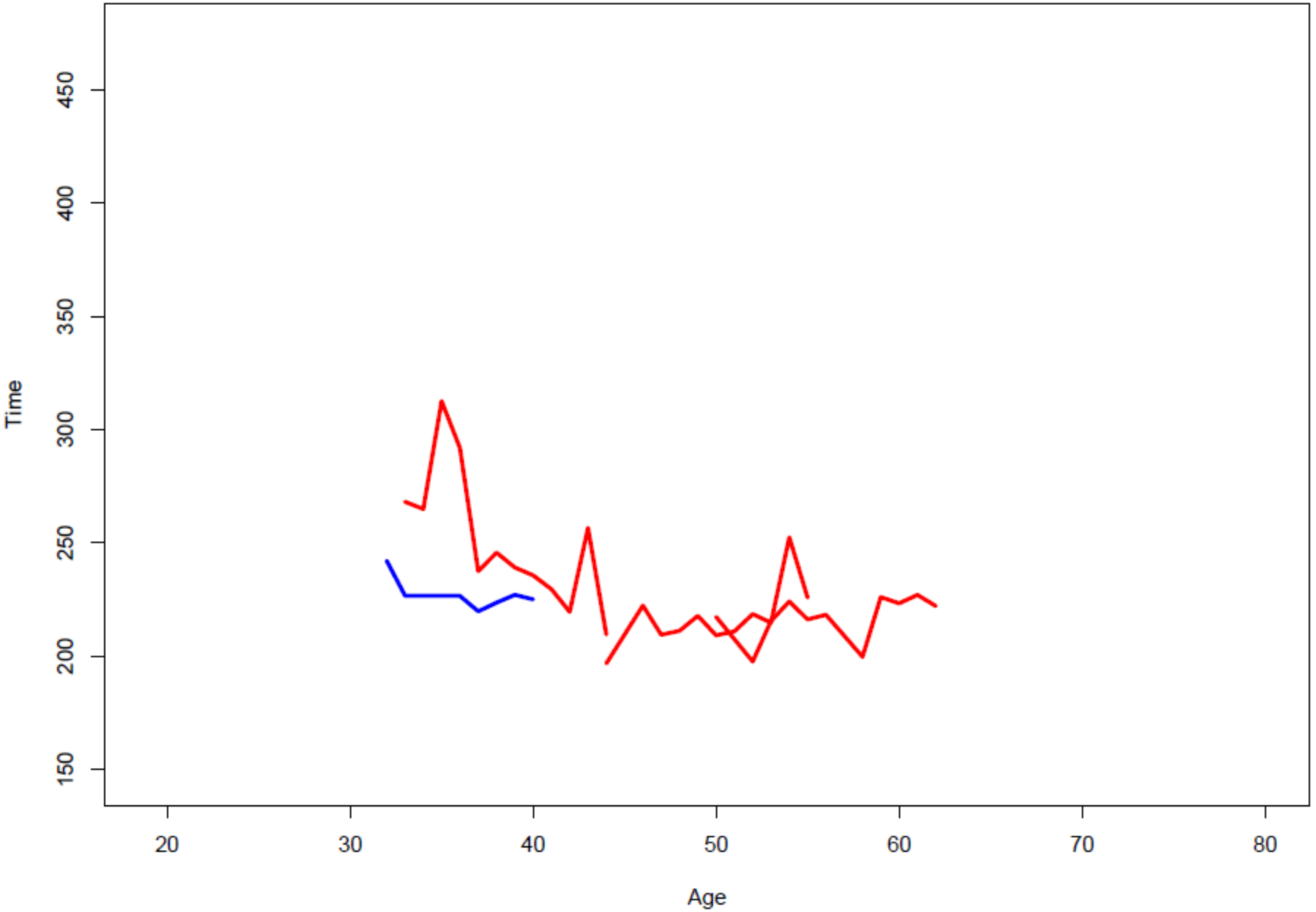
# *Part II: Longitudinal Approach to the Performance v. Age Problem*

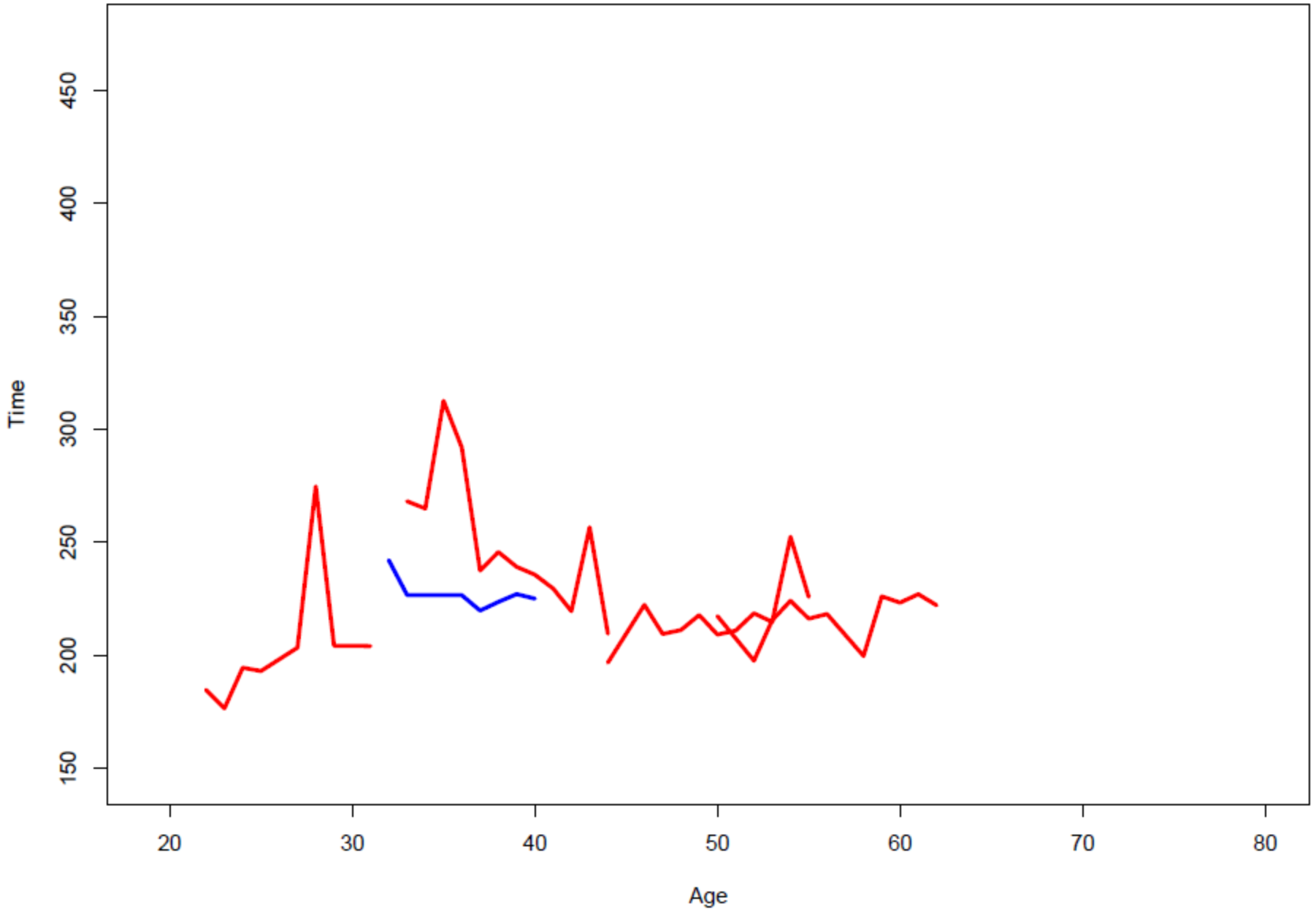
- About 500 runners have run the Boston marathon at least 10 years in succession (BAA)
- No easy way to reconstruct the list
- I used the datasets I had available to find 1,272 runners who had run each of 2010, 2011, 2013
- I then used the BAA archive to find all of those runners who had run at least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

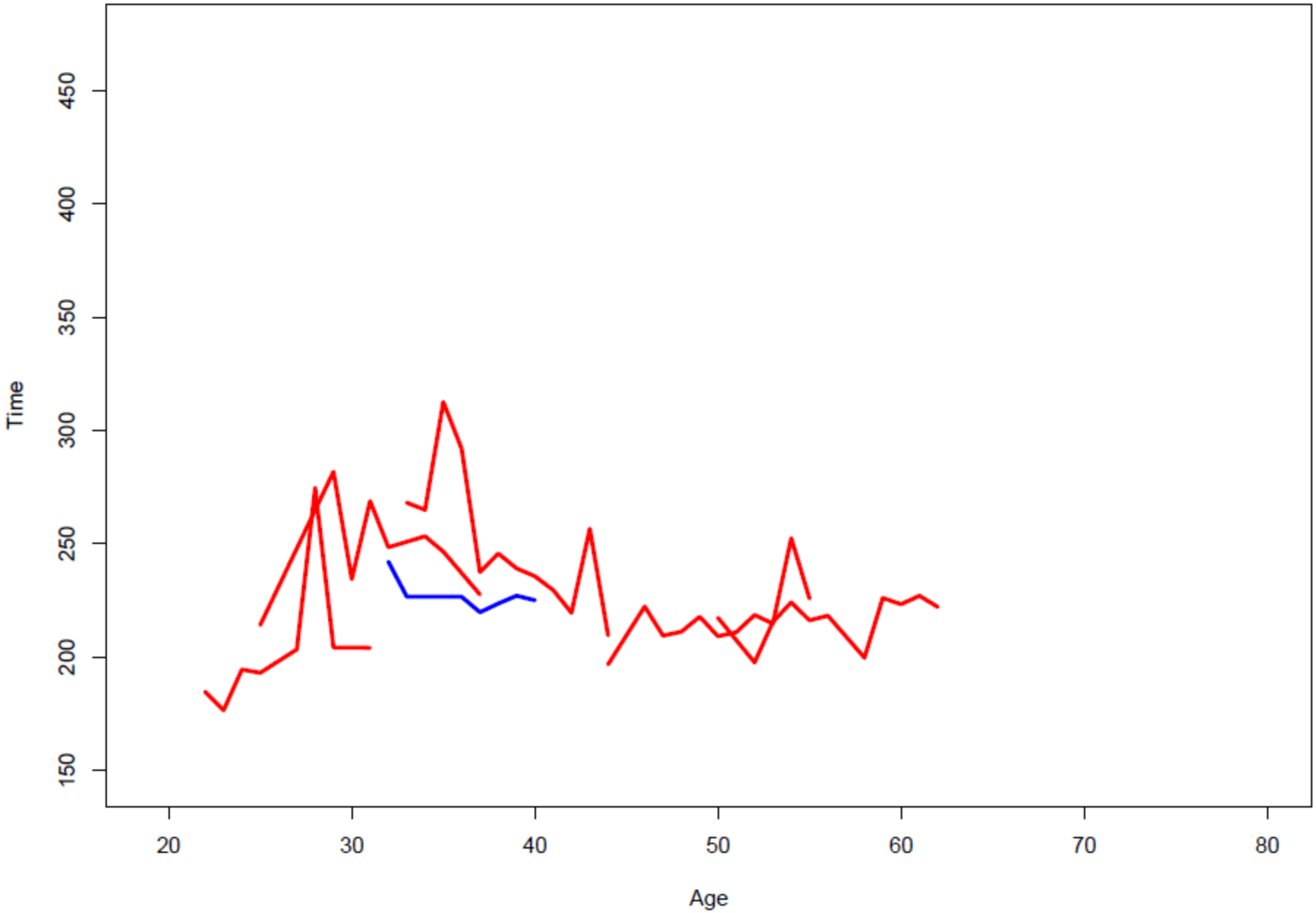

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

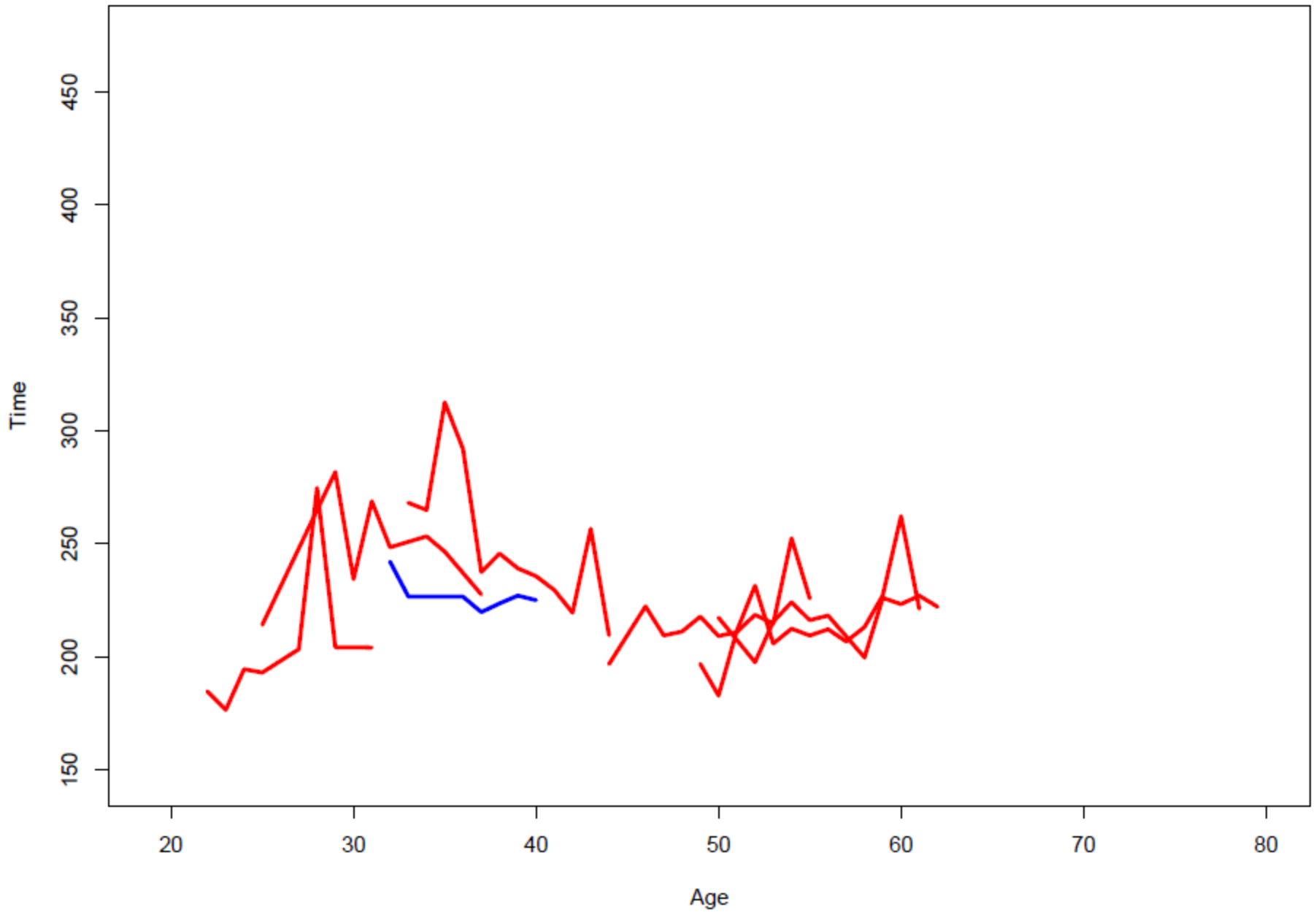

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

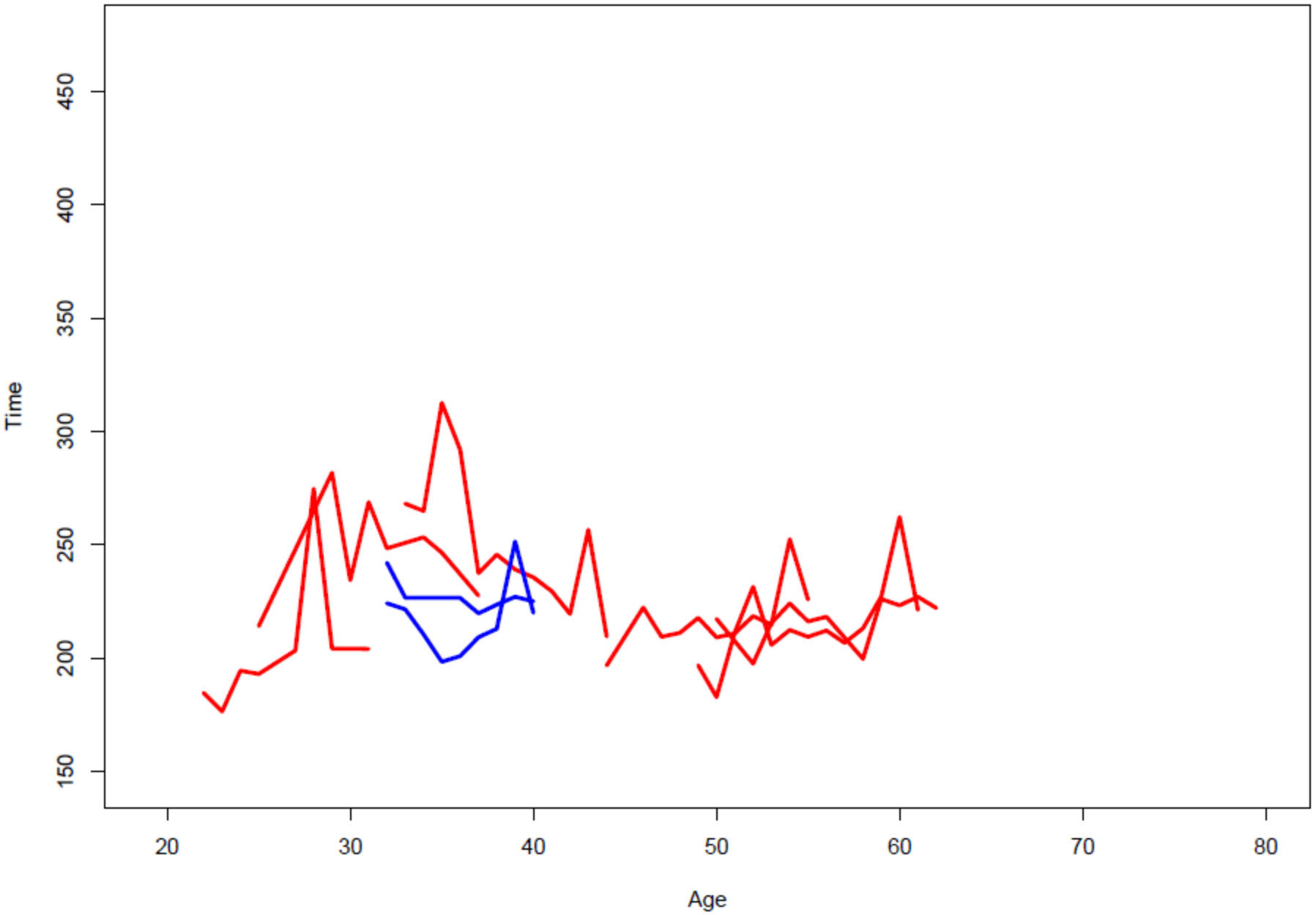

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

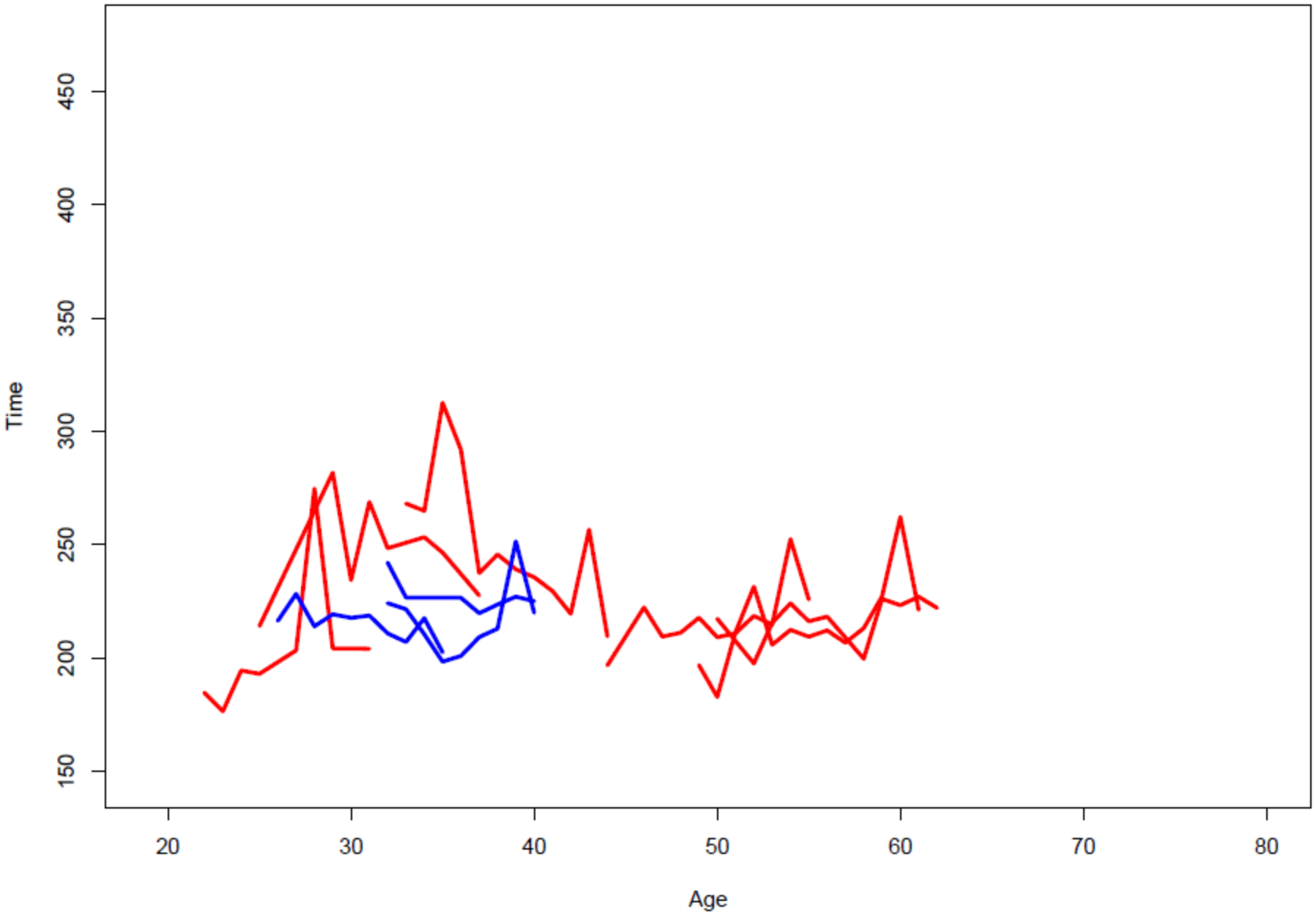

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

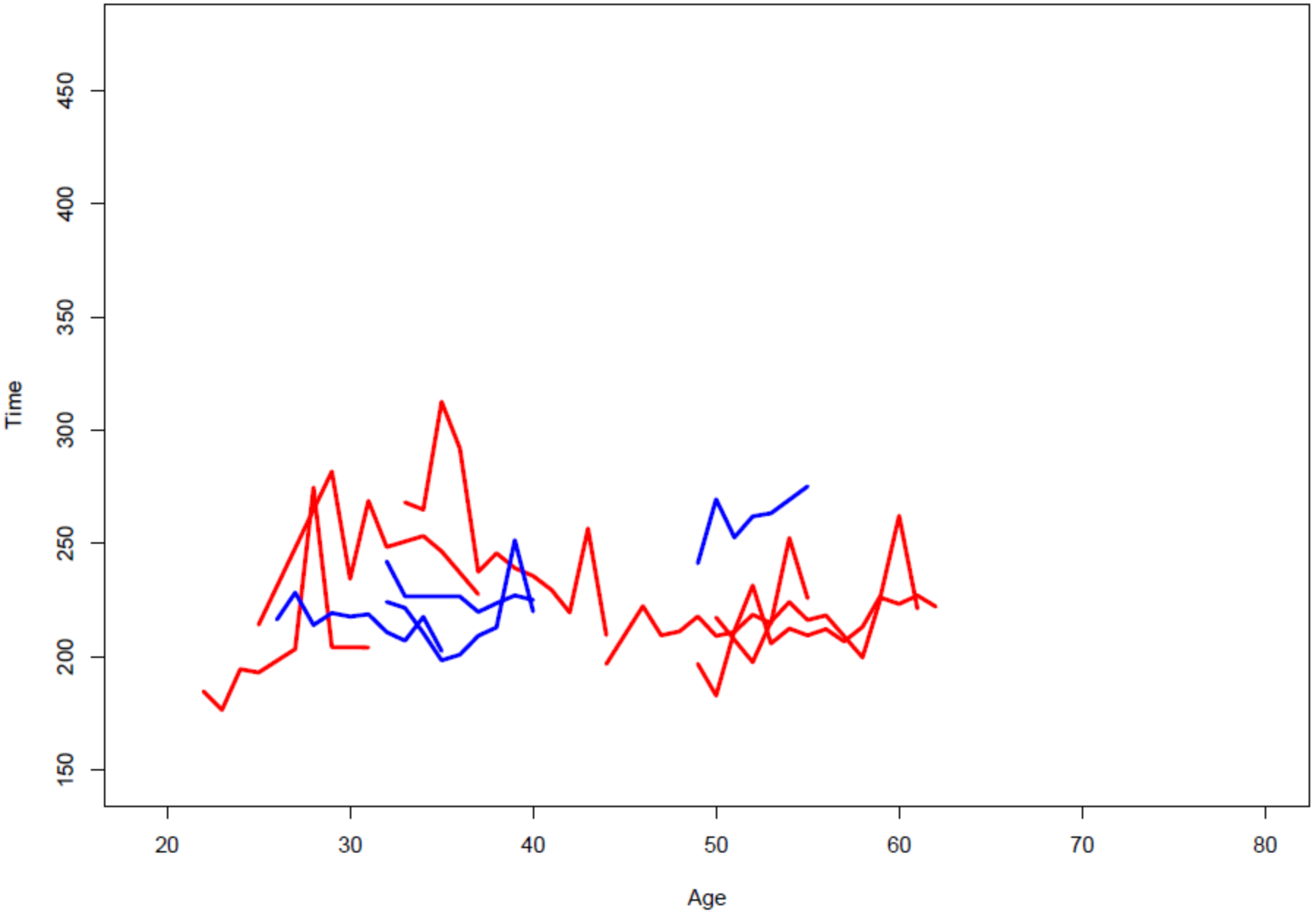

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

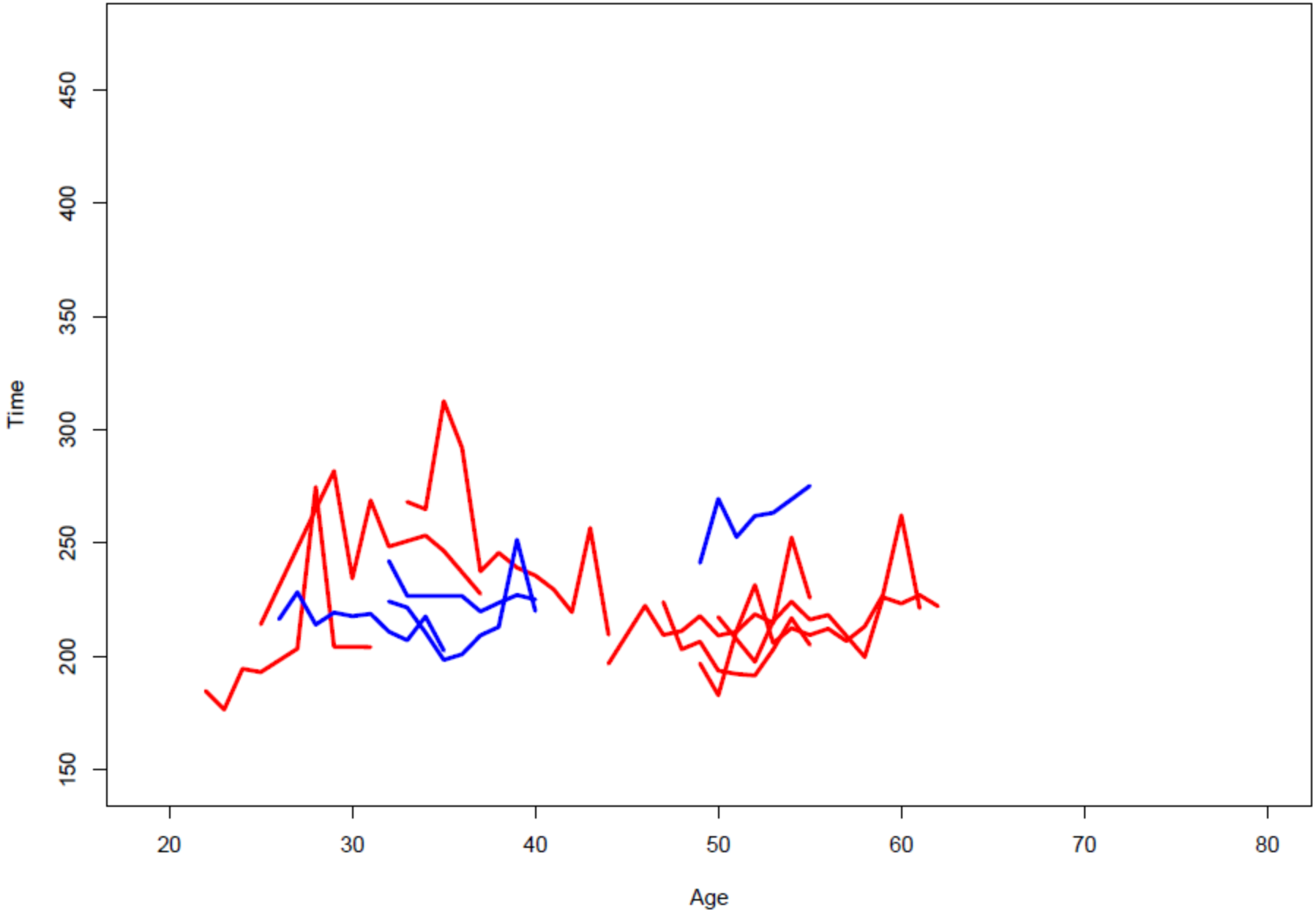

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

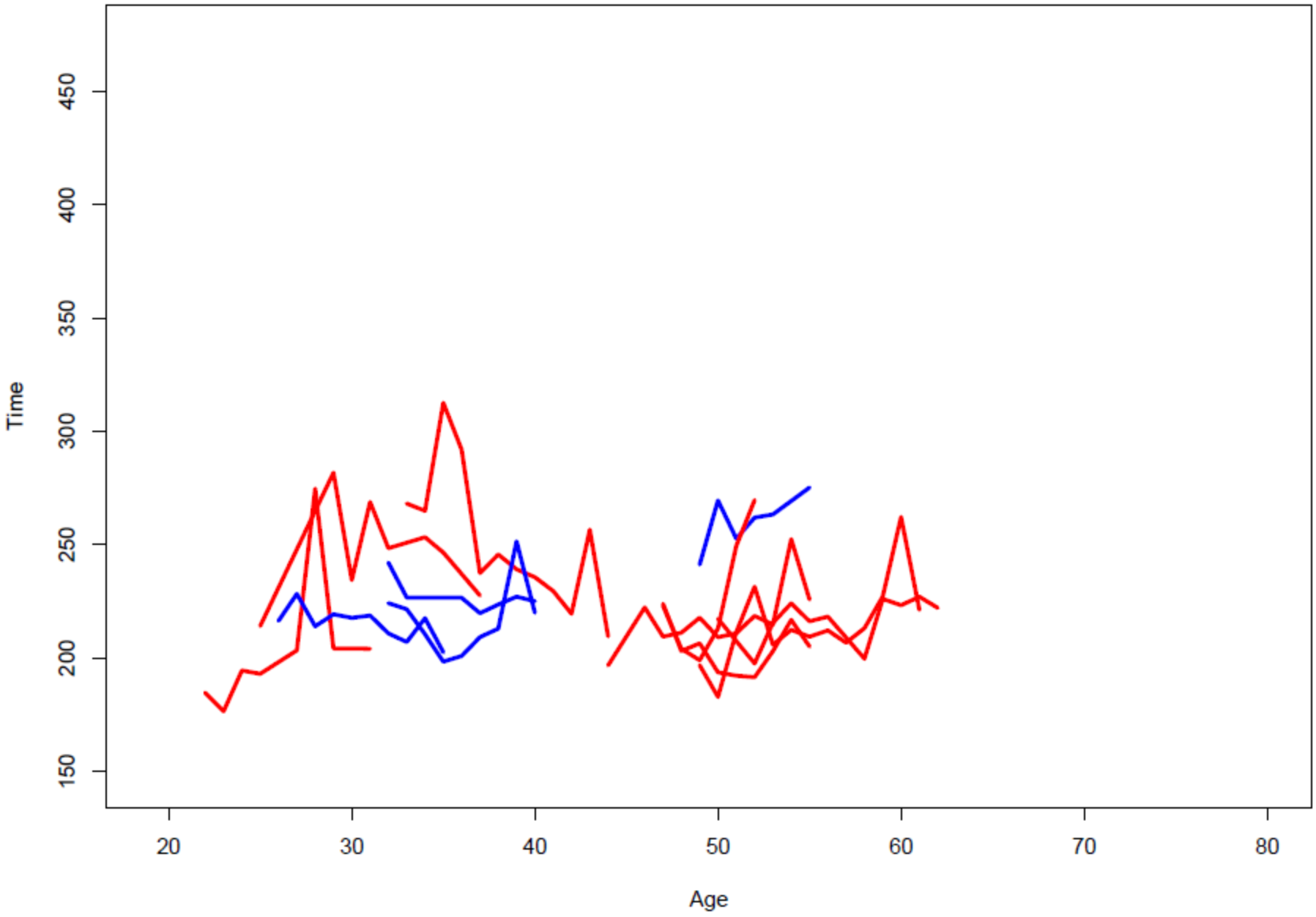

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

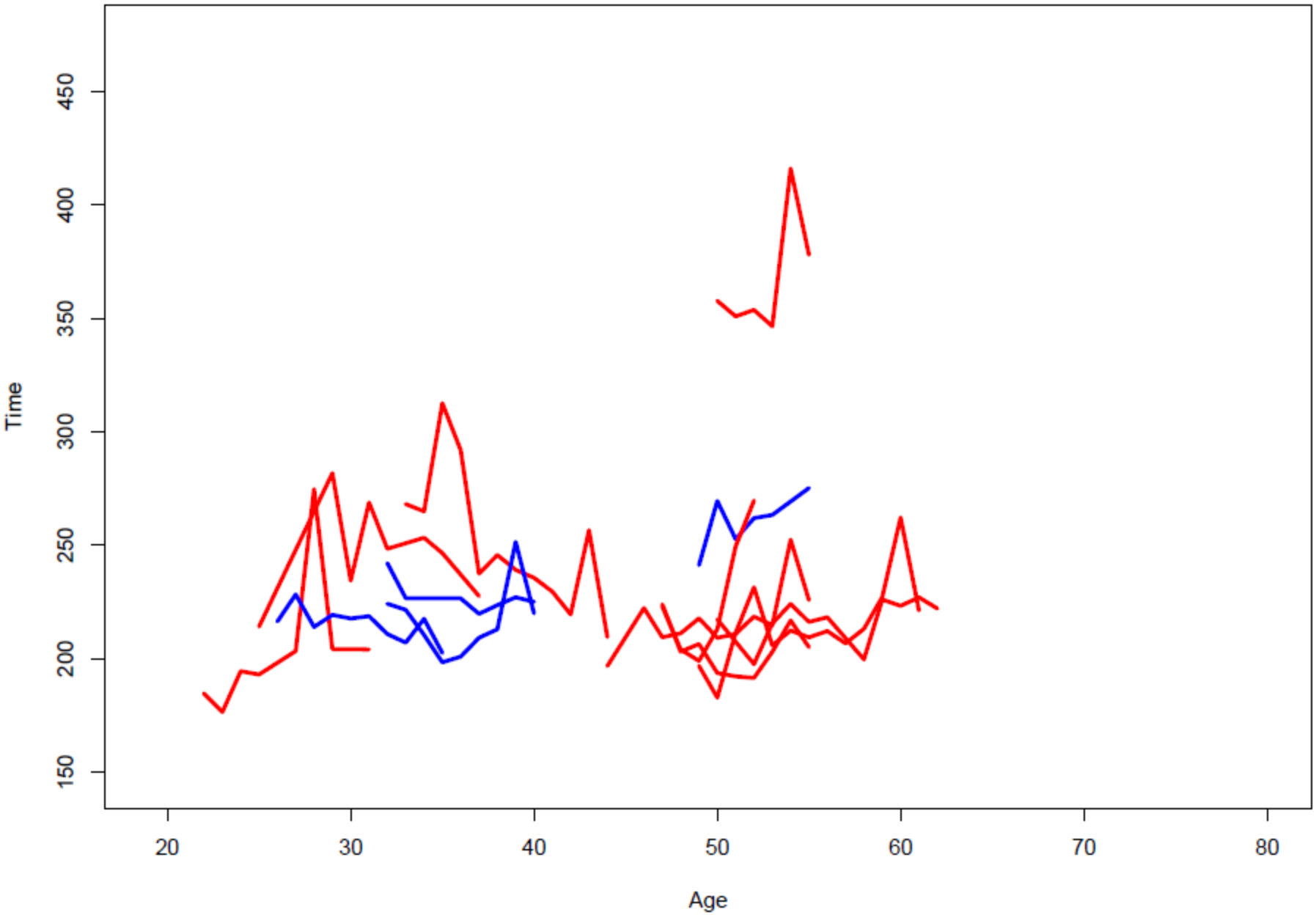

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

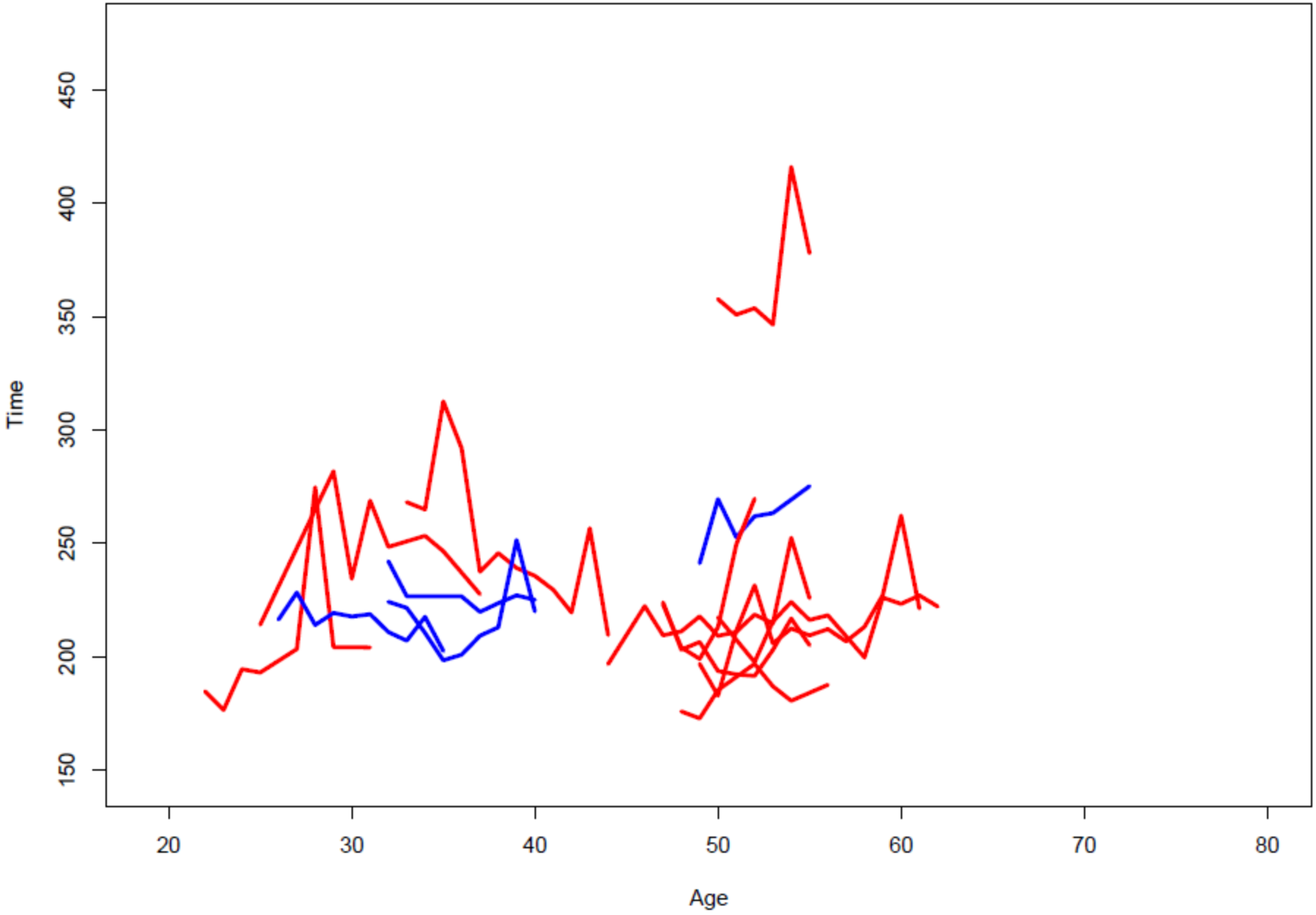

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

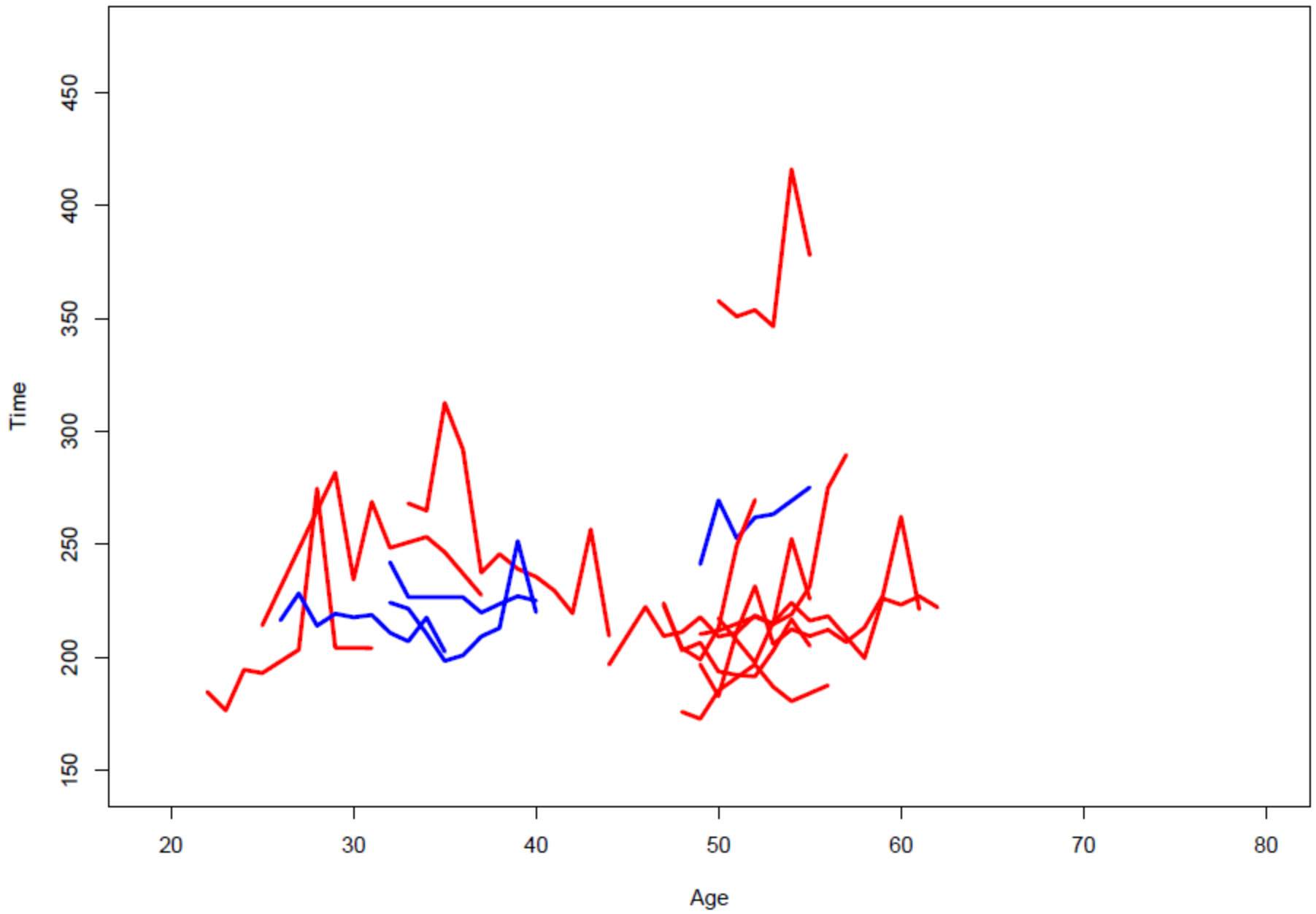

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

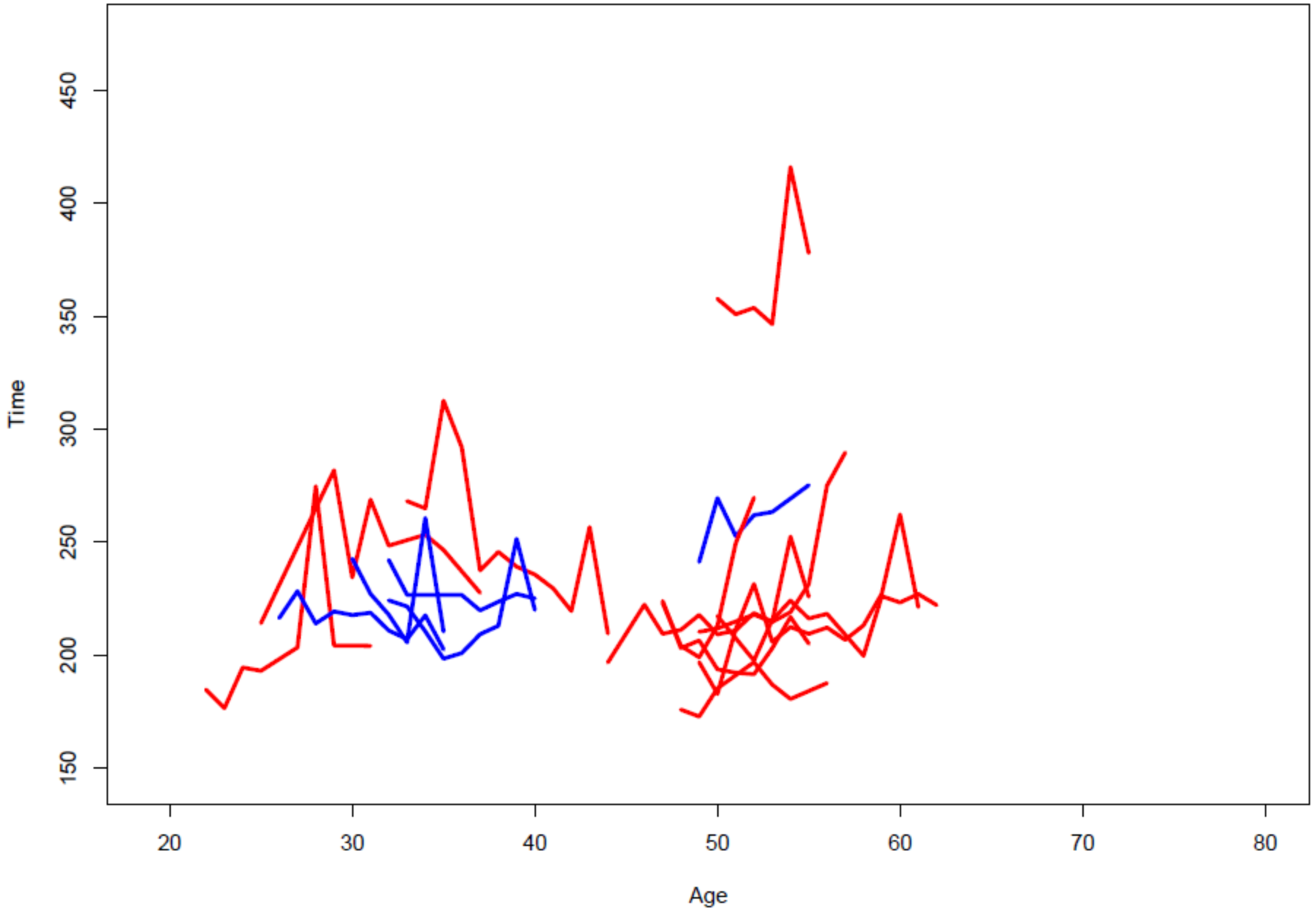

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

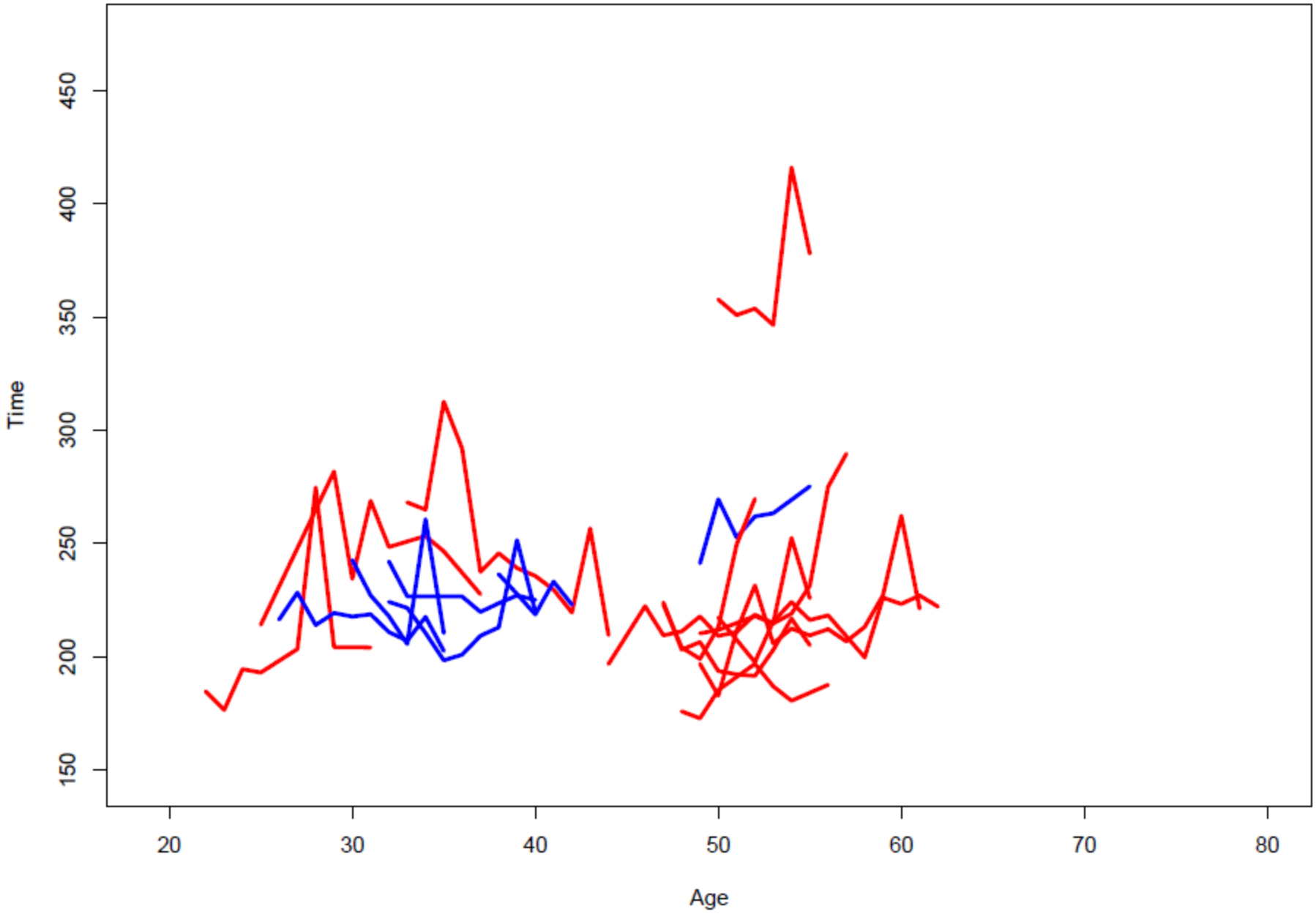

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

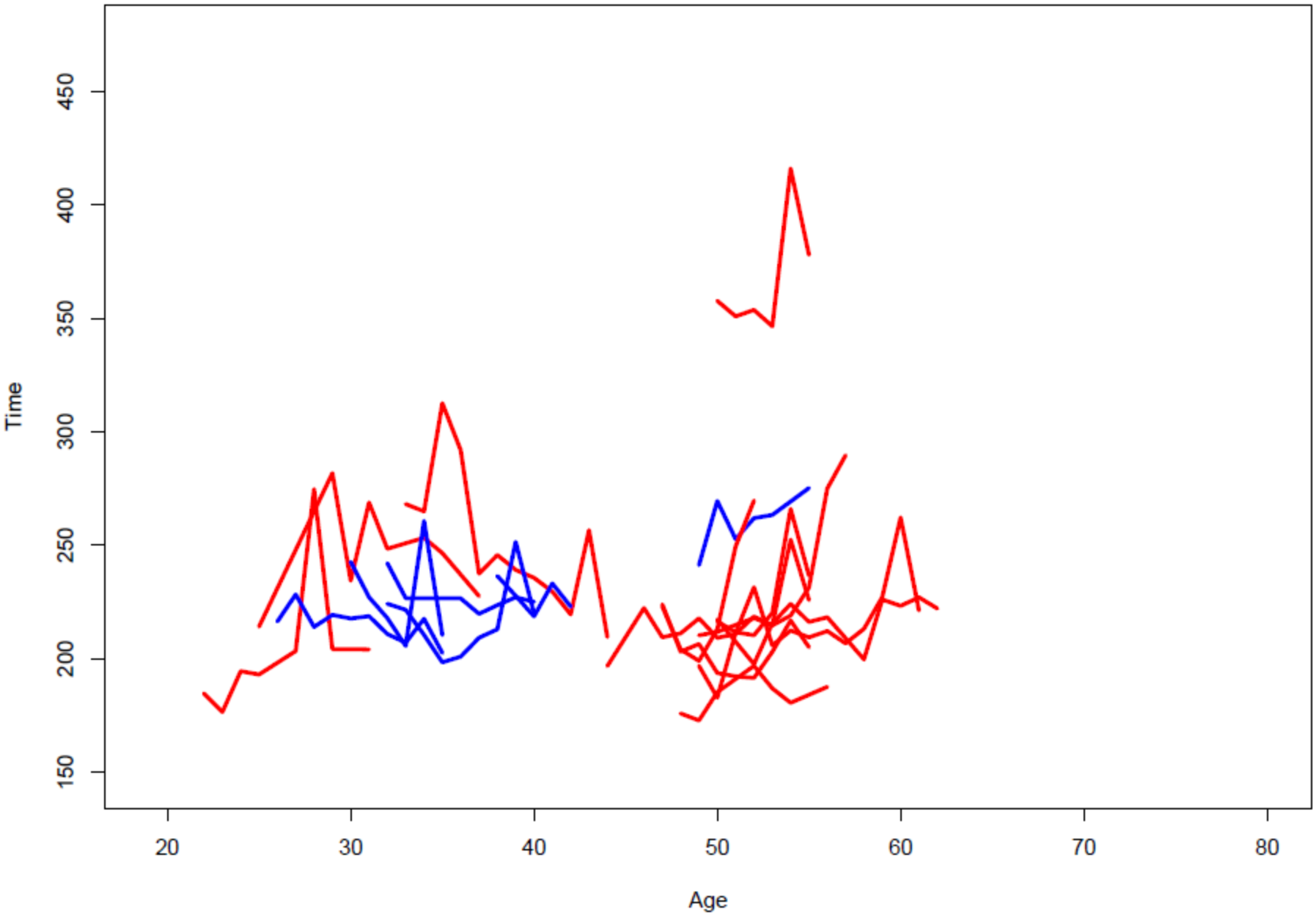

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

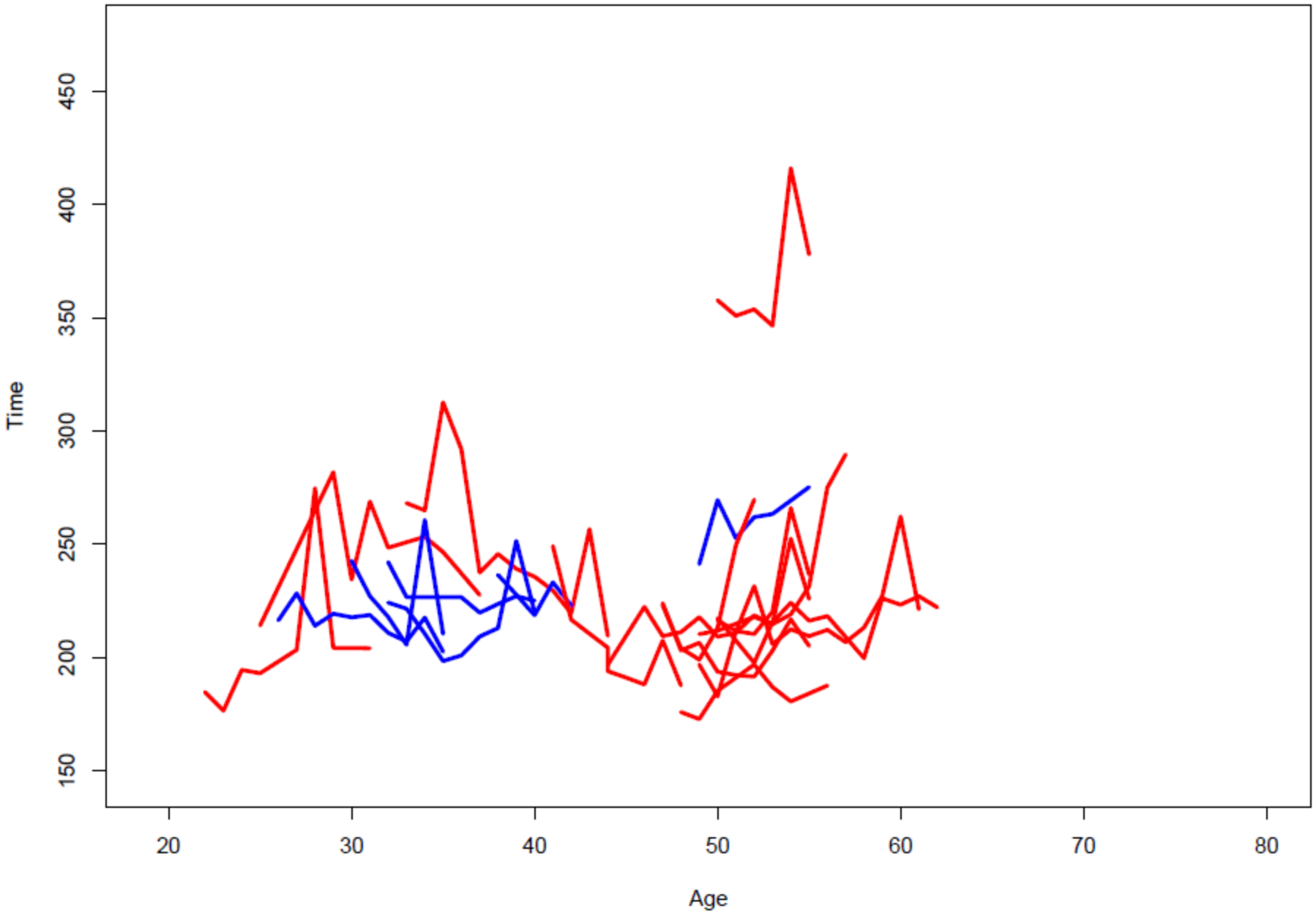

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

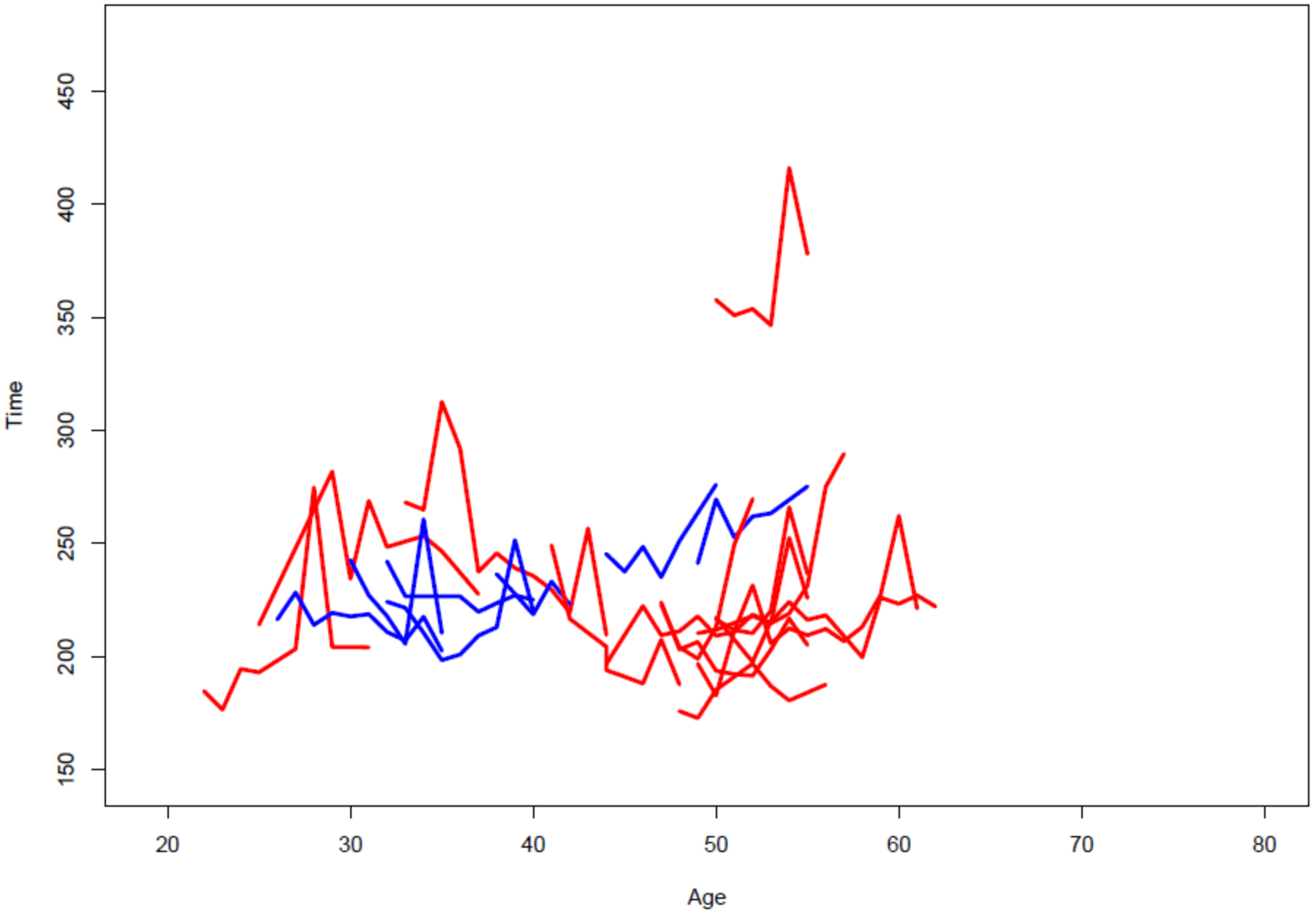

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age

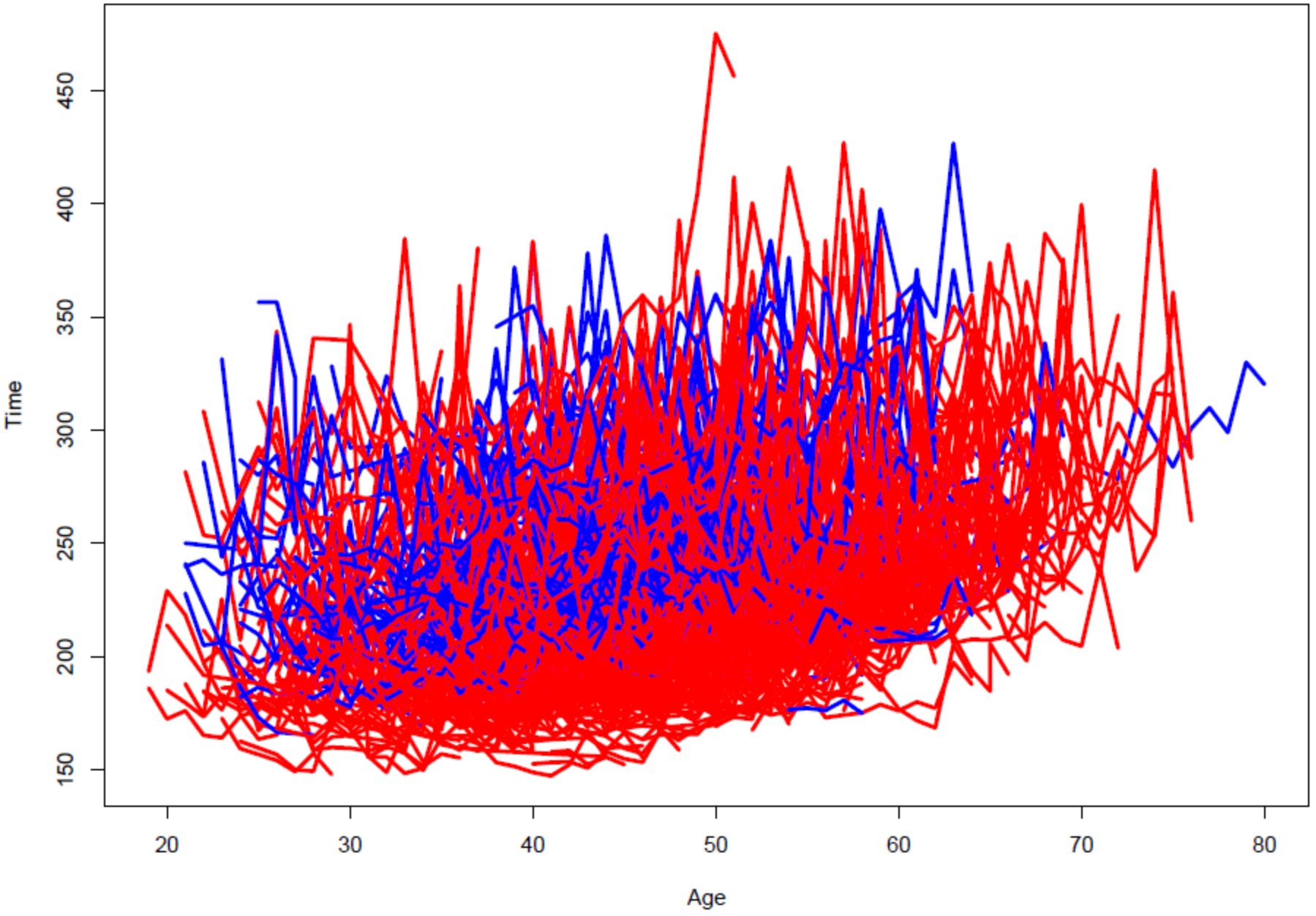

Time
150
200
250
300
350
400
450
20
30
40
50
60
70
80
Age


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*

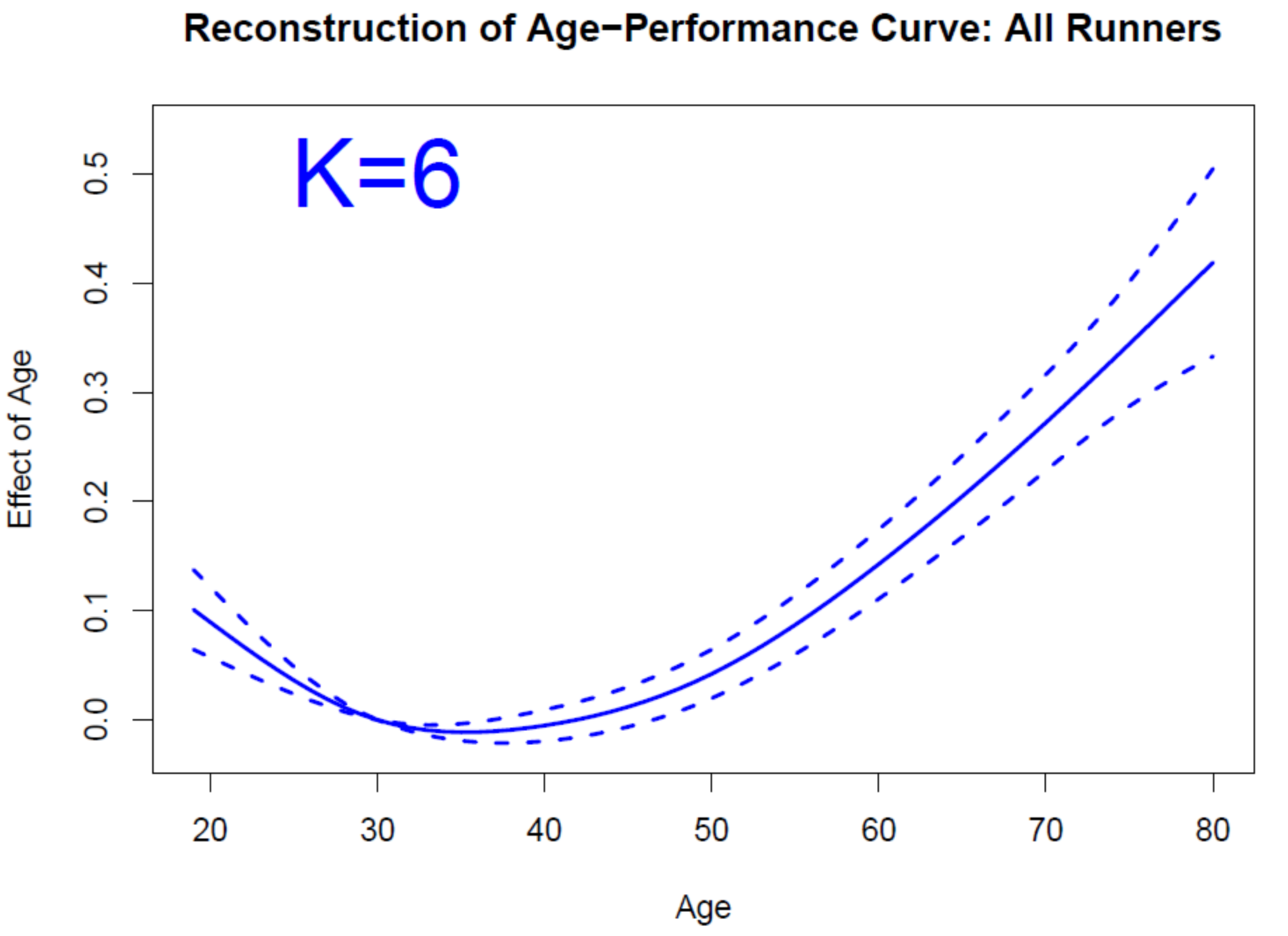
- Each individual runner record is a part-trace of the performance v. age curve for that runner
- Allow for a random “runner effect”
- Also allow for a random “calendar year” effect (2004 and 2012 were very hot)
- Separate men’s and women’s performance
- A refinement (later): also distinguish runners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

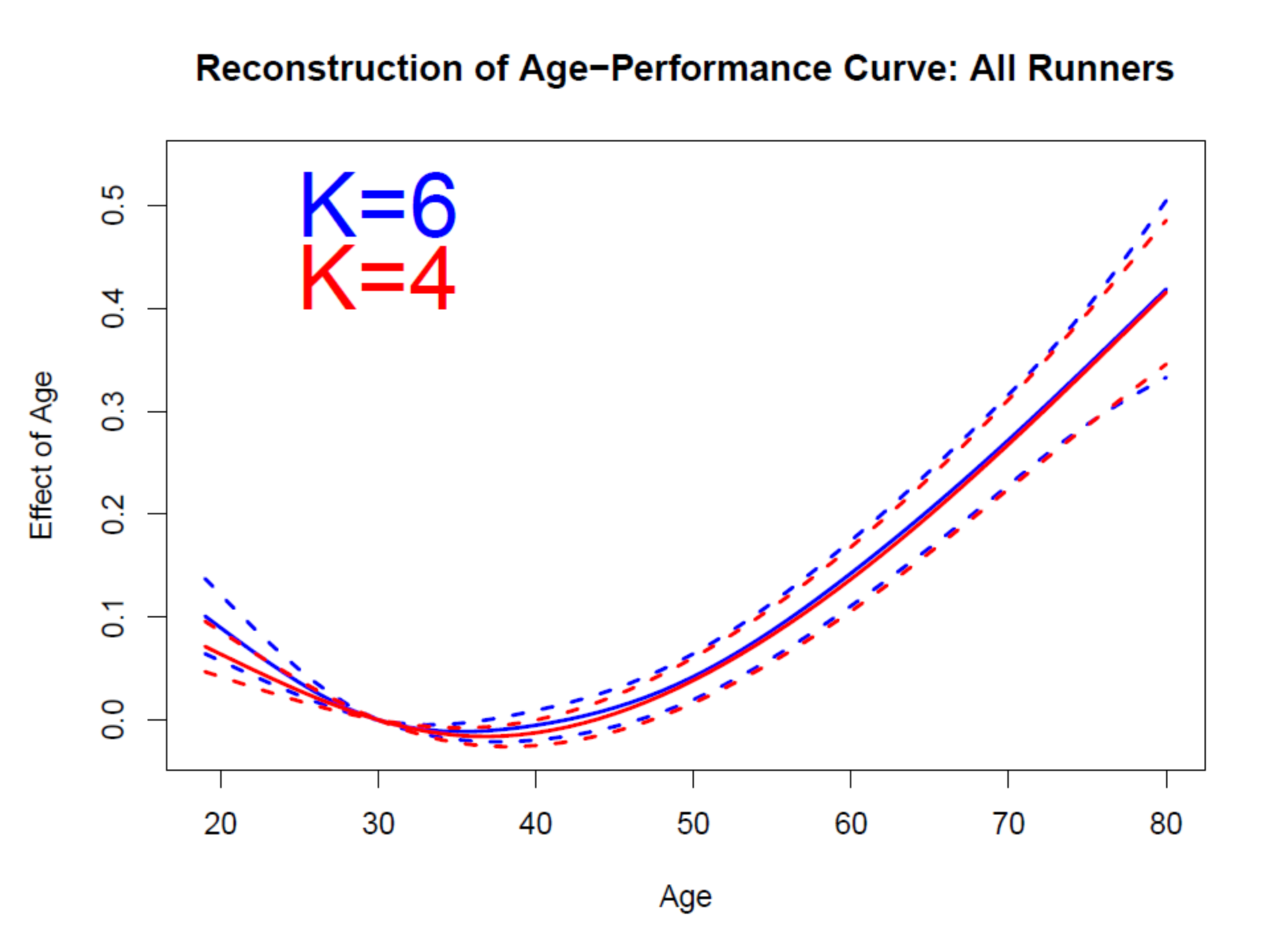
# Statistical Model:

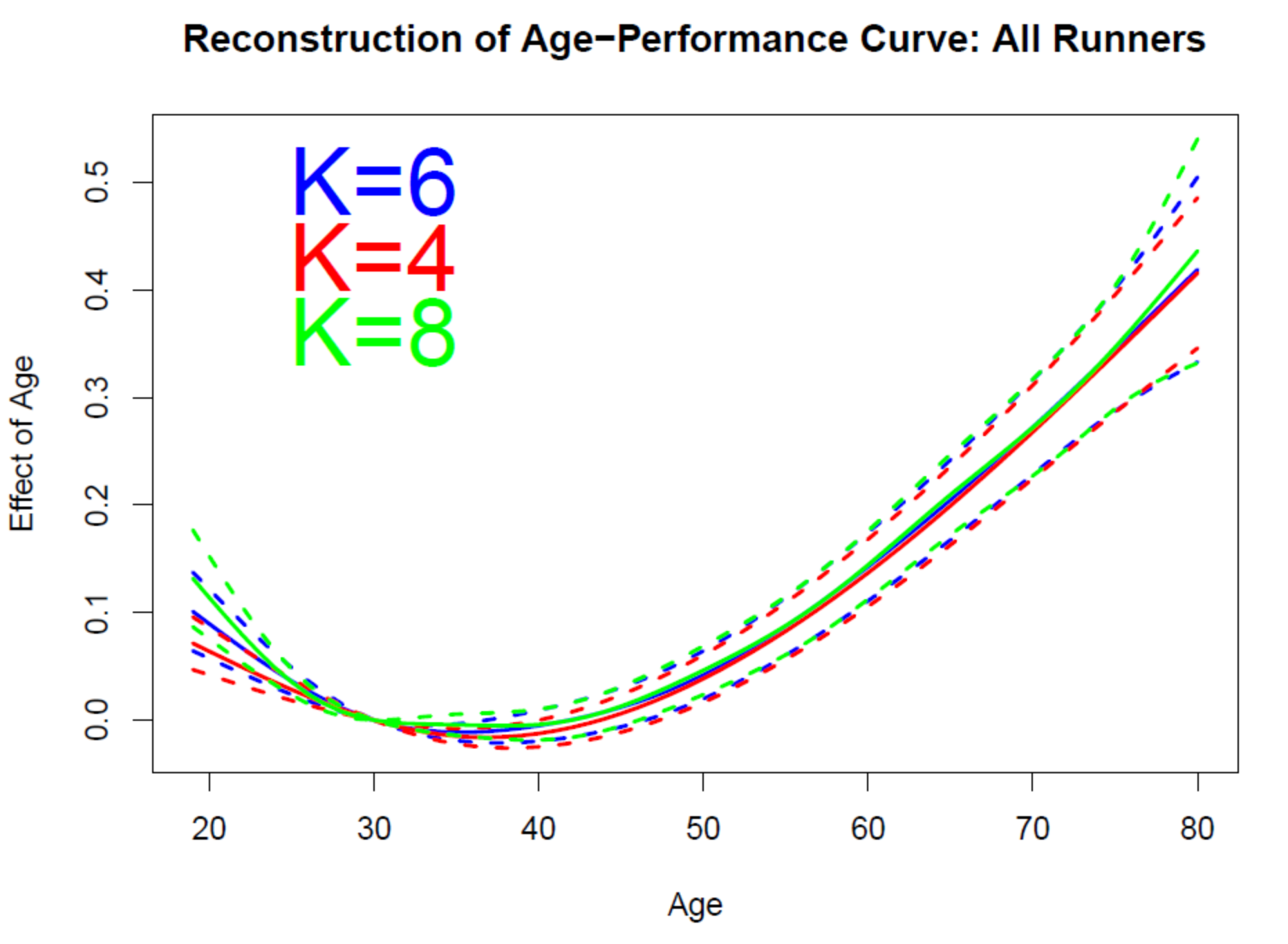
$$\log t_{ij} = \alpha_i + \beta_{y_{ij}} + S(a_{ij}; K) + \epsilon_{ij}, \quad (1)$$

where

- $t_{ij}$ is the $j$th finish time of runner $i$,
- $y_{ij}$ is the year of the $j$th finish time of runner $i$,
- $a_{ij}$ is the $i$th runner's age in her $j$th finish time,
- $\alpha_i$ represents the overall ability level of runner $i$ (small $\alpha_i$ 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,
- $\epsilon_{ij}$ is a random error.
  - In practice replace $S(a_{ij}; K)$ with $S(a_{ij}; K) - S(30; K)$.
  - Use R package `lme4` to fit.

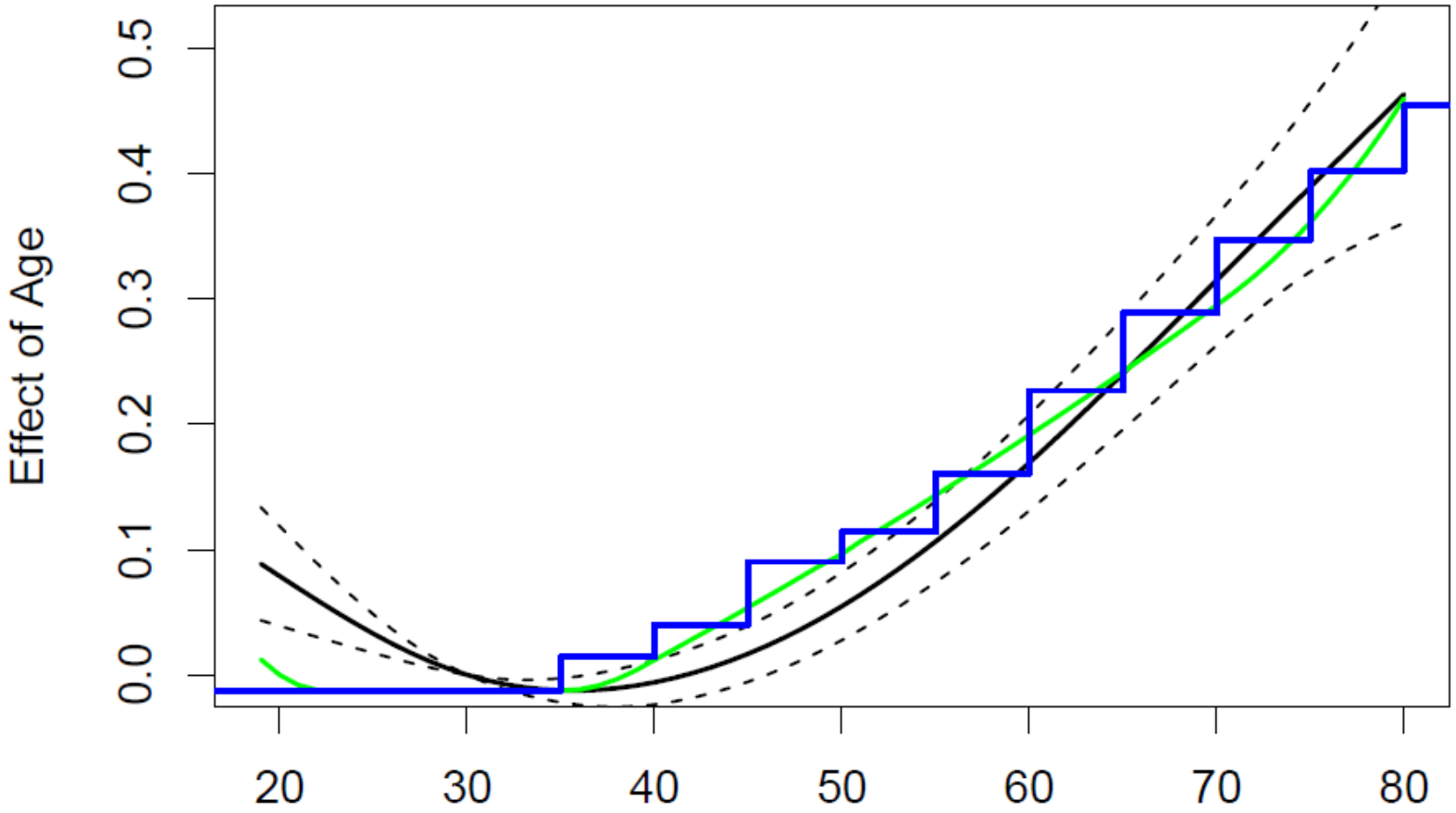

Reconstruction of Age−Performance Curve: All Runners
K=6
Effect of Age
0.0
0.1
0.2
0.3
0.4
0.5
20
30
40
50
60
70
80
Age


Reconstruction of Age−Performance Curve: All Runners
K=6
K=4
Effect of Age
0.0
0.1
0.2
0.3
0.4
0.5
20
30
40
50
60
70
80
Age

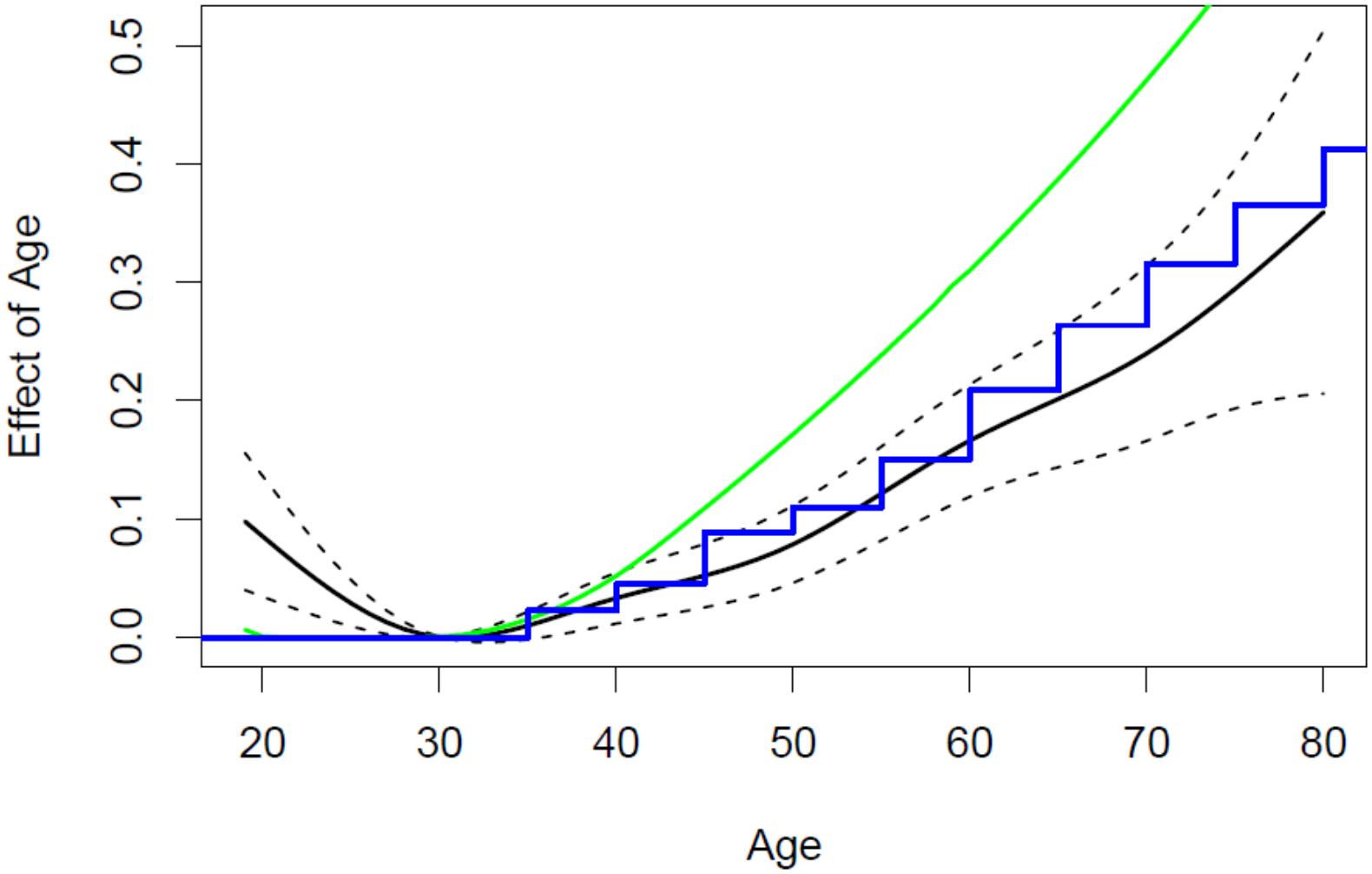

Reconstruction of Age−Performance Curve: All Runners
K=6
K=4
K=8
Effect of Age
0.0
0.1
0.2
0.3
0.4
0.5
20
30
40
50
60
70
80
Age

In practice, we compute separate curves for men and women and also compare with the age-performance curves that are implicit in the current Boston Marathon qualifying times, and in age-graded performances

Age-graded performances: a method of comparing performances for different age/sex categories by relating them to current world or national records

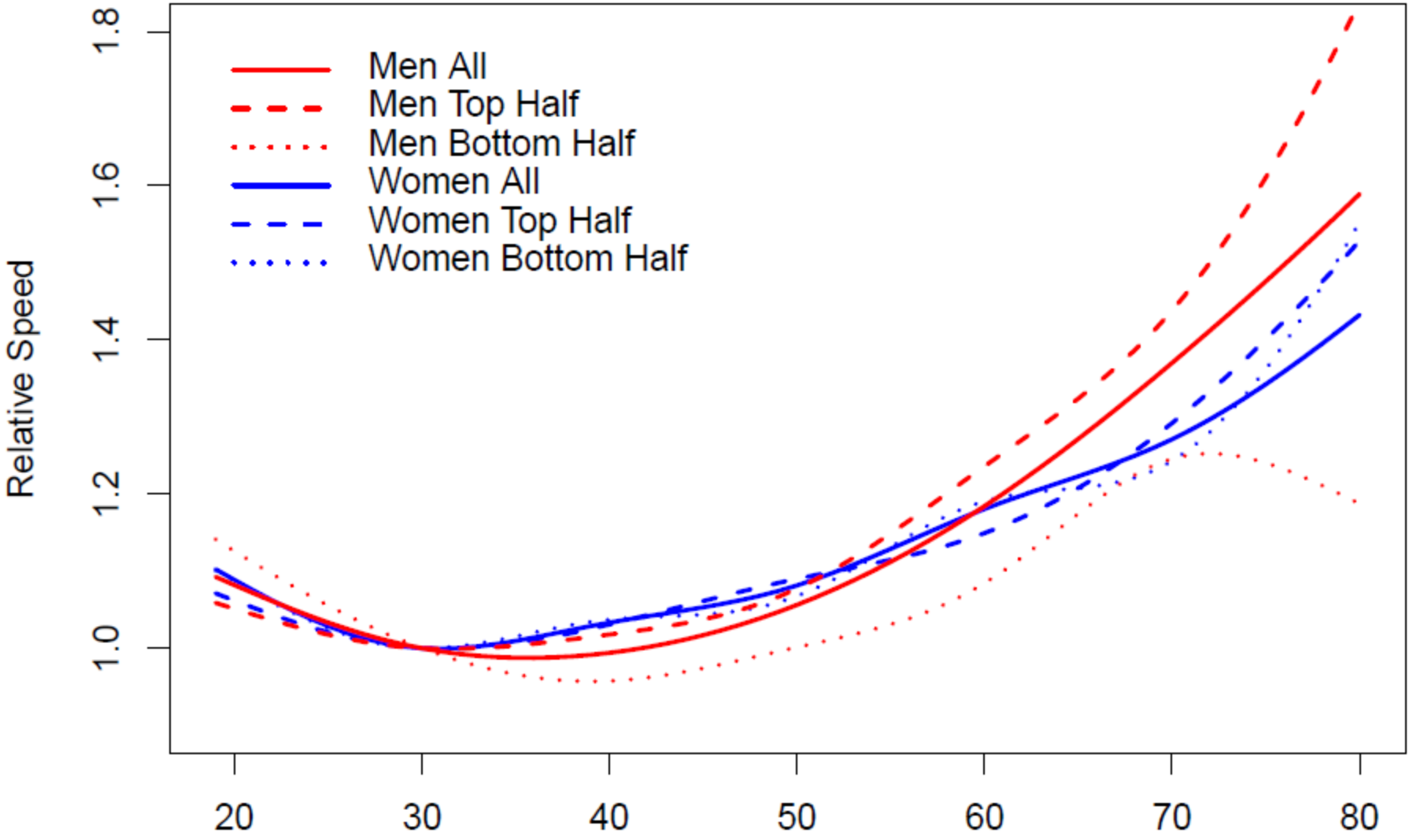



Age-performance curve for men. Solid black curve is spline-based estimate of age effect relative to age 30; dotted black curves are pointwise 95% confidence bounds. Blue step function based on Boston marathon qualifying times; green curve based on age-graded performances. Both the blue step function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

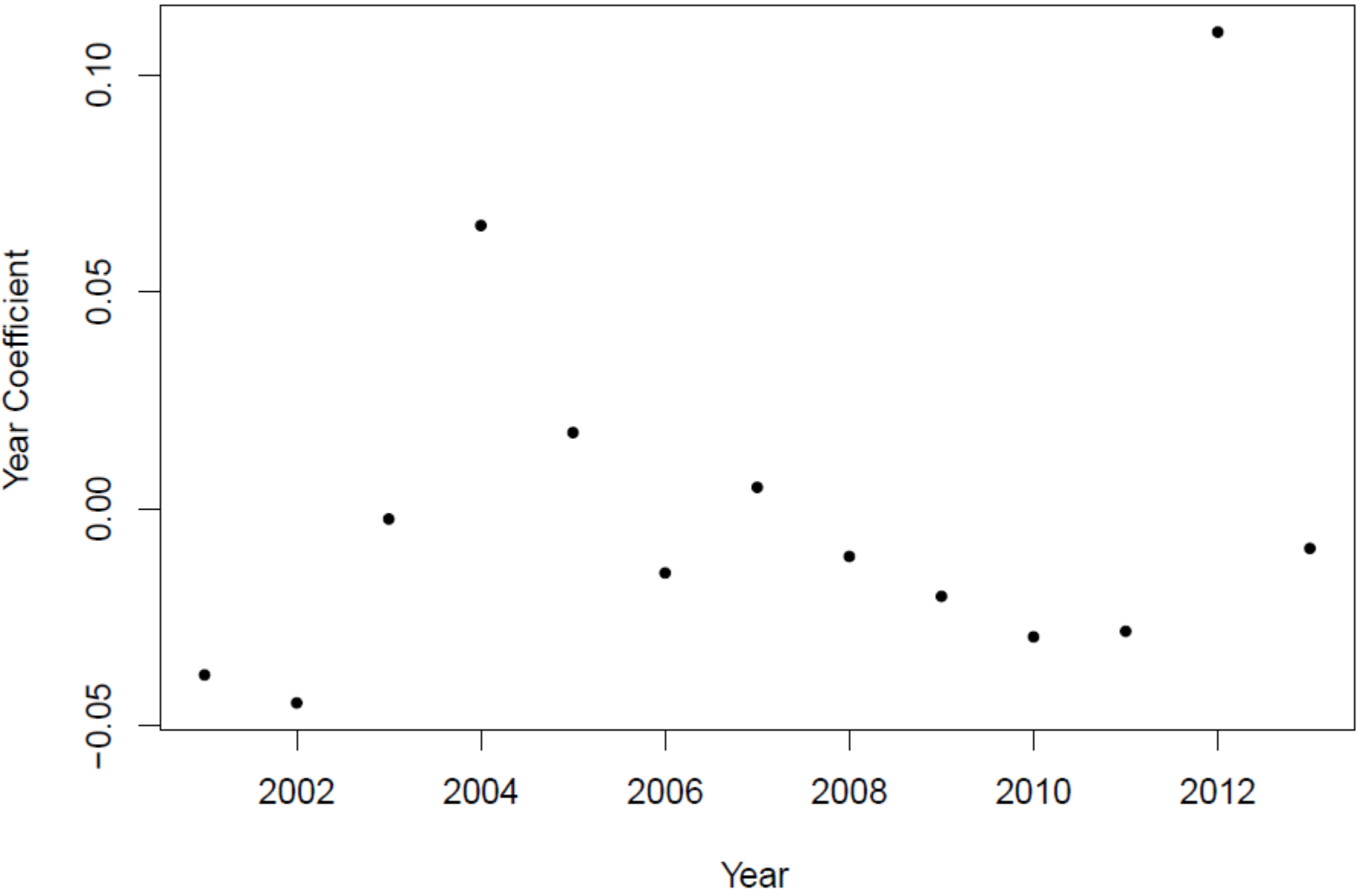



Same as previous page, but for women's performances.

# *We can also directly compare men with women, and different subsets of each*



# *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
- Still need larger datasets and more development of the method

# *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
- Still need larger datasets and more development of the method
- Thank you for your attention!