CrossMark

ARE WOMEN AS COMPETITIVE AS MEN IN ROAD RACES?

Jennifer Johnson[1] --- Rhonda Magel[2][†]
[1,2]Department of Statistics North Dakota State University Fargo, ND

ABSTRACT

Samples of road races are taken for various distances, including 5K, 10K, half-marathon, and marathon. Age/gender adjusted times are calculated for each runner of every road race included in the samples. Hypothesis tests are conducted comparing the average age/gender adjusted running time for men with the age-gender adjusted time for women for each race in the sample and then using meta-analysis techniques to collectively compare the mean adjusted times for each of the various distances. Hypothesis tests are also conducted comparing the variances of the age/gender adjusted running times for women with the variance of the age/gender adjusted running times for men, as well as the proportions of men and women who have age/gender adjusted running times greater than the 60th percentile, and then less than the 40th percentile. Overall, the average age/gender adjusted running times for women are less than those of men for each of the distances. It is found that a significantly higher proportion of men perform at or above the 60th percentile compared to the proportion of women for each of the race distances. It is found, however, that the variances of the age/gender adjusted running times for men are larger.

Keywords: Age/gender adjusted running times, Means, Variances, Proportions at or above the 60th percentile, Various distances, Meta-analysis.

Contribution/ Originality

This study contributes to the literature by comparing the proportions of male and female runners who finish at or above their 60th percentiles of age/gender adjusted running times and who finish at or below their 40th percentiles of age/gender adjusted running times for each of the major road race distances.

1. INTRODUCTION

Running is easily accessible to any person who is interested in trying it. This is just one of the many aspects of running that has contributed to the massive increase in the popularity of the sport. The total number of finishers for road races in the United States has been increasing greatly over since 1990, from 4.8 million to 19.03 million in 2013 (Running USA, 2014). Along with the increase in numbers, the years have also seen an increase in the proportion of female race finishers, from twenty-five percent in 1990 to fifty-seven percent in 2013 (Running USA, 2014). This increase in popularity has brought up questions regarding the differences between male and female athletes, specifically, whether the differences between the populations of finishers substantiate the beliefs that are held by the public.

The common belief among the public is that men are more competitive than women, especially when it comes to sports. When asked about what kind of runner they would classify themselves as, more women chose to identify as a fun runner or a fitness participant while more men chose to identify as a competitor (Running USA, 2014). These results appear to back up the public's perception of male and female runners, but a direct comparison of race

results disagrees. Between 1976 and 2005, the best male marathon performance dropped 2.83%, but the best female marathon performance dropped 14.46% (Burfoot Amby, 2007). The difference between these performances also changed notably from an 18.71% difference in 1976 to a 7.73% difference in 2005 (Burfoot Amby, 2007). The divergence between the views of the public, the attitudes of the runners, and the actual results needs to be examined thoroughly.

When the performances of men and the performances of women are only compared within their respective genders the gap between male and female competitiveness can only be estimated. In order to determine the truth behind the opinions of the public, we need to be able to compare the performances between genders. In 1989 the World Masters Athletics (WMA), formerly the World Association of Veteran Athletes (WAVA), developed the first set of tables for grading the performance of an athlete based on sex and age which were updated, with the help of Howard Grubb and Alan Jones, in 1994, 2004, and most recently in 2010 in order to more accurately estimate the performance data (World Master's Athletics, 2010). These age grading tables were designed to enable runners to compare performances to all other running performances without worrying about a gender or age bias. In this paper, we will analyze the results, using age/gender graded scores, of multiple races in an attempt to determine if any differences are present in the data. To calculate the age/gender graded scores for this paper, we used the 2010 open standard tables, the finish times of the racers, and the ages of the runners and inserted these values into the

following equation: $AG = 100 * \frac{Open\ Standard\ for\ Distance}{Age\ Factor*Time(in\ seconds)}$ (WMA, 2010).

The underlying motivation for our research is the wide held belief that men are more competitive than women and, based on this belief, we formed several hypothesis to test using our research data. If men are truly more competitive than women, we would expect to find significant differences between the two populations when compared statistically. In respect to age/gender graded scores, we would like to determine if men finished comparatively faster than women by looking at proportion of runners that finish with an age/gender graded score of sixty percent or higher and the proportion that finished with an age/gender graded score of seventy percent or higher. We would also like to investigate if women finish comparatively slower than men by looking at the proportion of runners that finish with an age/gender graded score of forty percent or lower. We believe that male race finishers have a higher average age/gender graded score than female race finishers overall; we also believe that females are more variable, in age/gender graded scores, than male race finishers.

Our specific research hypotheses are formally written out below:

- Age graded scores of sixty percent or higher and seventy percent or higher

H_0: Proportion of Male Finishers = Proportion of Female Finishers

H_α: Proportion of Male Finishers > Proportion of Female Finishers

- Age graded scores of forty percent or lower

H_0: Proportion of Male Finishers = Proportion of Female Finishers

H_α: Proportion of Male Finishers < Proportion of Female Finishers

- Are women more variable than men?

H_0: Variance of Male Finishers = Variance of Female Finishers

H_α: Variance of Male Finishers < Variance of Female Finishers

- Mean age/gender graded scores

H_0: Mean Age/Gender Graded Score of Male Finishers = Mean Age/Gender Graded Score of Female Finishers

H_α: Mean Age/Gender Graded Score of Male Finishers > Mean Age/Gender Graded Score of Female Finishers

- Median age graded scores

H_0: Median Age/Gender Graded Score of Males = Median Age/Gender Graded Score of Females

H_α:Median Age/Gender Graded Score of Males > Median Age/Gender Graded Score of Females

The actual age distributions of men and women runners will also be examined.

2. REVIEW OF LITERATURE

There is no current literature available that concerns our exact topic, but there is a plethora of literature concerning the underlying topics of our research – the competition between men and women and the history of marathon running for women.

We will discuss the competition between men and women. This competition can be broken down into several distinct forms: competition in athletics, competition in everyday life, and the reasons behind the competition. The topics are often intermingled when reviewed and researched in the literature. Researchers often use a setting of competition in order to test the beliefs that are held by male and female contestants, both consciously and unconsciously. Two papers that follow this design were written by Dato and Nieken (2014) and Booth and Nolen (2011).

Dato and Nieken (2014) examined the behavior of males and females in a controlled tournament setting with the opportunity for sabotage. They noted that although the females make up nearly half of the workforce, they are underrepresented in the upper levels of companies worldwide. The study was designed to examine the reasons behind this disparity by implementing four different treatments: baseline, belief, cheating, and gender. The researchers found that males, on average, selected twice as much sabotage as females, but that the genders did not differ significantly on any other treatment investigated. Dato and Nieken (2014) concluded that males and females systematically differed in their sabotage decisions, resulting in males winning the tournament more often, and that the gender gap was already present in the beliefs of the participants before they entered the study.

Booth and Nolen (2011) also examined the behavior of men and women in a controlled competition setting to determine why women hesitate to compete while men choose to compete too much. They theorized that if men enjoy competition and women dislike it there will be two effects: fewer women will choose to compete and fewer women will succeed in competitions. In order to test these theories, the researchers designed a controlled experiment constructed around boys and girls solving puzzles in three rounds: solving puzzles in groups, mandatory tournament, and choice of whether to enter the tournament or not. The children also filled out an exit survey so that researchers could determine the background information and beliefs of each participant. Booth and Nolen found that girls solved fewer puzzles than boys in all settings and that the environment mattered in the decision to enter the tournament. The results from the study supported the following beliefs: men choose to enter competition more than women, girls in same gender groups enter competition more than girls in mixed gender groups, and girls from same sex schools enter competition more than girls from coed schools. Specifically, single sex schooling decreases the gender gap in choosing to compete by over sixty percent; girls from same sex schools are forty-two percent more likely to enter the tournament than girls from a coed school and girls from coed schools are seventy-one percent less likely to enter the tournament than boys from coed schools.

While researchers use competition to investigate how men and women react to competition and what their beliefs about competition are, surveys and discussion are used to study the reason behind why men and women compete in the first place. We reviewed two papers that follow this study design, the first was by Warner and Dixon (2015) and the other was written by Cashdan (1998).

Women are often viewed as not aggressive or competitive enough to succeed in the sports setting. Warner and Dixon (2015) theorized that men and women tend to view competition differently in addition to possessing physiological, psychological, and social reactions to it. The researchers conducted interviews, both individually and in groups, with current and former collegiate athletes to determine how competition impacted their sport experience. The interviews focused on internal and external competition as well as the reasons behind the competitive feelings. Both genders viewed external competition as enhancing the sport experience by making it

more meaningful as there was something at stake for the participants. There were differences in respect to internal competition; women found that it was often taken to a personal level while men affirmed that it created and nurtured a mutual respect among the competitors. Both genders agreed that when teammates were motivated by trying to beat each other out of a spot on the team, rather than bettering themselves, then competition could become problematic. The interviews suggested that, while competitiveness may be a biological tendency of males, the social arena surrounding sports creates an atmosphere where men feel the need to live up to an expectation of being competitive and where women feel the need to restrain from it.

Cashdan (1998) conducted a behavioral study using diary entries to investigate the reasons behind competition between men and women. The participants were asked to fill out forms about when they felt competition, their background information, and how they felt about feeling competitive. These forms were all filled out by college students in a school setting, which may have influenced the results. Cashdan found that students were more likely to report conflict which resolved in their favor and that more forms focused on school related conflicts than sports related ones. Additionally, the few forms that were related to sport conflicts were written by men.

Each of these papers, approaches the topic of male and female competition in a different manner and from a different angle. In contrast to the reviews of the reasons behind competition and the reactions to competition, the papers comparing the differences between the performance of male and female runners generally focus on the gap between the sexes. Several papers that studied this gap in performance are by Frick (2011); Thibault (2010) and Pate and O'Neill (2007).

Frick (2011) used the top two hundred finishes from professional distance running at the four most popular distances: five kilometer (track), ten kilometer, half marathon, and marathon; only including races where men and women were competing against their own sex. The differences between the genders were calculated as a percentage of male finishing time, for both genders. Frick found that culture and incentives, rather than predisposition and biology, are responsible for the competition. Thibault (2010) also compared same sex events between male and female athletes and calculated the gender gap based on the male results. Specifically, the researchers compared the world records for eighty-two Olympic events and found that gap between male and female world records has been stable since 1983. Thibault et al. also found that the number of women participating in the Olympics has increased at a higher rate than the number of men since 1986, while the frequency at which the world records have been set for women have decreased in that time. These results could also be as a result of the increase in the number of Olympic events available to women. In the 1984 summer Olympic Games, 221 total events were available with only 76 offered for women; by contrast, the 2012 summer Olympic Games hosted 302 total events with 144 of those events available for women (International Olympic Committee, 2015).

Pate and O'Neill (2007) looked specifically at the marathon time differences between male and female runners between 1976 and 2005. Before 1976, the public believed that women were not physiologically able to run a marathon and were barred from running long distances. Despite the obstructions, women's participation has grown dramatically over the years. The researchers found that from 1976 to 2005 the marathon times of American men stayed fairly constant, from 2:10:10 to 2:11:14, while the times of American women decreased from 2:47:10 to 2:21:25, a difference of 15.6%. Finally, we will examine the literature surrounding the history of women's running. Women's distance running came into the public viewpoint during the 1960s and entered the Olympic Games in 1984. The popularity of women's distance running has increased since the marathon first appeared in the 1984 Olympic Games in Los Angeles with the total number of finishers for the marathon and half marathon increasing every year. Between the years of 1990 and 2013, the total number of marathon finishers increased from 224,000 (10% female finishers) to 550,637 (43% female finishers) and the total number of half marathon finishers increased from 303,000 to 2,046,600 (61% female finishers) (Running USA, 2014).

The men's marathon was first contested in the Olympics in 1896 followed by the 5,000 meter (five kilometer) and the 10,000 meter (ten kilometer) races in 1912 but the women's marathon was not contested until 1984

followed by the five kilometer race in 1988 and the ten kilometer race in 1996 (International Association of Athletics Federations, 2015). The majority of the literature concerning the history of women's distance running consists of discussion about the fight for the inclusion of the marathon to the women's Olympic track and field program. There are many sources that review the fight for the addition of the women's long distance running in the Olympic Games and we reviewed two for this paper: Lovett (1997) and Schultz (2015). Each of these sources took a slightly different approach to the data but both covered the main points of the struggle for inclusion, detailed below.

Women were excluded from track and field competition entirely until the 1928 Olympics where they were granted the opportunity to run the 100 and 800 meter races. The media reported that many of the female competitors had difficulties finishing the 800 meter race due to poor preparation and that multiple women collapsed upon finishing the race but the historical records show that all nine contestants finished the 800m finals without a single woman collapsing on or after the finish line. Nevertheless, the International Olympic Committee (IOC) built on the false reports and removed the 800 meter race from the women's program, considering it too strenuous for women. The IOC restricted women's track and field races to 100 meters and under until 1948 when the 200 meter race was added to the roster. The women's length restrictions were increased again in 1960 when the 800 meter race was reintroduced and in 1972 when the 1500 meter race was added but the distances were not extended again until 1984 when the 3000 meter and marathon races were added (Lovett, 1997; Schultz, 2015).

Prior to 1967, very few women ran in major competitions because, outside of the Olympic Games, women were banned from taking part in organized distance running races. In 1967, rules forbade men and women from racing in the same event as well as prohibiting women from competing in races over 1.5 miles (2.4 kilometers). That same year K. Switzer was able to obtain an official registration to the Boston marathon and, after race officials attempted to forcibly remove her from the course, women's long distance running was thrust into the public eye around the world. This global attention, in addition to the performances of early women marathoners, elicited an increase in the popularity of women's distance running in addition to triggering a gradual transformation of the American and international rules for distance running (Lovett, 1997; Schultz, 2015).

Throughout the 1970s the worldwide popularity of women's marathon running exploded, from around twenty in 1970 to more than eight thousand in 1979. This expansion brought about changes to the guidelines concerning long distance running for women, by the end of the decade the Amateur Athletic Union (AAU), the International Association of Athletics Federations (IAAF), and individual events around the world officially allowed women to compete in marathons. Despite the clear popularity of the marathon among female competitors, the IOC did not include the race in the 1980 Olympic Games. The IOC stated several qualifications that needed to be met before an event could be added to the Olympics, the most difficult being that the event needed to be popular in at least twenty-five countries and on at least two continents. The popularity of the women's marathon came to the attention of the international press when the third Avon international marathon championship was covered by major worldwide media in 1980, as a result of the boycott of the Olympic Games in Moscow (Lovett, 1997; Schultz, 2015).

Women also needed to show the IOC that they could maintain an organization and race structure, separate from men. To accomplish this many companies, led by Avon Inc. and Nike, began to support the women's running movement by creating women's only races, producing ad campaigns that featured women running, and putting pressure on the IOC to include the marathon, along with the 5k and 10k race distances, in the 1984 Olympic Games. These advocacy groups were put to the test when the IOC and the Los Angeles Olympic Organizing Committee (LAOOC) attempted to delay making the decision for inclusion of the women's marathon in the 1984 Games. The LAOOC initially attempted to block the inclusion stating that expanding the women's program would create an excessive strain on the Olympic officials and the city of Los Angeles; this verdict was quickly reversed after the United States Olympic Committee, the IAAF, and several LA bureaucrats received heavy criticism from women's advocacy groups. The IOC had cited the need for more medical research data on the effects of long distance running on the female body and the need for more experience for women. The support groups, and the

popularity of the third Avon International marathon championship, put pressure on the IOC by publicly refuting these excuses, given as an attempt to delay making a decision on the inclusion of the women's races, using medical professionals and the research that had been done to that point (Lovett, 1997; Schultz, 2015).

Faced with all of the data, the IOC voted to include the women's marathon and the 3000 meter race in the 1984 Olympic Games and the LAOOC quickly accepted the decision. The women's running movement, and their supporters, succeeded in their mission to add the marathon but the mission to add the 10,000 meter and the 5,000 meter races would not succeed for another four years for the 10,000 meter race and another twelve years for the 5,000 meter race (Lovett, 1997; Schultz, 2015).

2.1. Data Description

Our overall data is comprised of the age and gender of the runner, the timed race results, and the associated age/gender graded scores in terms of proportions, calculated using the WMA age-grading tables developed by Howard Grubb, with one dataset for each major road running distance (Five kilometer, ten kilometer, half marathon, and marathon) and a separate dataset for major marathons. Our goal is to examine any differences that exist in the race populations and if those differences could have an effect on the age grading calculations. This paper will focus on two separate areas of analysis: Qualifying races and non-qualifying races. Qualifying races were defined as marathons that require a runner to reach a qualifying time or participate in a specific qualifying race prior to registration.

For the marathons which require runners to qualify before participating, referred to as major marathons, we will be including races from 2013 and 2012 depending on the race. The 2012 running of the New York marathon was cancelled in connection with the destruction of the Superstorm Sandy hitting the east coast (Runner's World, 2013) and the 2013 running of the Boston marathon was disrupted when two bombs exploded near the finish line (Runner's World, 2013). As a result, we included the results from the 2013 New York marathon and the results of the 2012 Boston marathon in our results. For the Chicago marathon, no particular difficulty occurred during either the 2012 or 2013 races so we included the 2013 results, in fitting with our previously stated plan. The Boston marathon and the New York marathon only allow runners aged eighteen years and older while the Chicago marathon allows runners over the age of sixteen. In order to be consistent across the races, we only included data from finishers that were eighteen years old and older in our research for these marathons.

To get a larger picture of the differences between male and female running populations, we will consider the results from non-qualifying races at each of the four popular running distances: marathon, half marathon, ten kilometer, and five kilometer. The number of finishers in road races varies greatly between race distances so we will select results from eight races of varying sizes for each distance. By selecting variable sized races we feel confident that the data set is a good representation of the US road running population. Healthy running practices observe that runners should limit competitive training and competition until they have undergone puberty which happens between the ages of thirteen and fifteen, depending on the gender of the athlete (Road Runners Club of America, 2015). In order to consider competitive runners and in an attempt to remove the possibility of counting a single runner twice in the same race, we only included runners aged fourteen years and older in our data. For each race, we will consider all of the age/gender graded scores and finishing times, as well as the ages and genders. All race results considered in this section were taken from races run during 2013.

3. RESULTS FOR PROPORTIONS- QUALIFYING RACES

For the first part of our analyses, we examined three marathons that require qualification before participation, hereafter referred to as major marathons. These races are three of the largest and most respected races in the United States: the Boston Marathon, the ING New York City Marathon, and the Bank of America Chicago Marathon. Specifically, we included data from the 2012 Boston Marathon, the 2013 ING New York City Marathon, and the 2013 Bank of America Chicago Marathon

All three races have over twenty thousand participants, with the exact numbers given in Table 1. The percentage of male and female race finishers with an age/gender graded score of sixty percent or higher and with an age/gender graded score of seventy percent or higher are given in Table 2. Examining each of the values in Table 1 reveals that the number of male finishers is larger than the number of female finishers for each race. From Table 2, we can see that the proportions of male finishers at 60% or higher (and 70% and higher) for each of the races is higher than the corresponding proportions for females. To test if these differences in proportions are statistically significant, we will compare the proportion of male finishers to the proportion of female finishers at each level of interest using a chi-square test of proportions (Weiss, 2012). For each of the three races, the proportion of male participants who finished with an age/gender graded score of sixty percent or higher is found to be significantly larger than the proportion of female participants who finished with an age/gender graded score of sixty percent or higher.

Table-1. Major Marathons by the Numbers

Marathon	Total	Male Finishers	Female Finishers
Boston	21,563	12,570	8,993
Chicago	38,757	21,416	17,341
New York	50,136	30,589	19,547

Source: (Boston Marathon, 2012; Chicago Marathon, 2013; New York Marathon, 2013)

Table-2. Proportion of Major Marathon Finishers and Test Results

Marathon		Proportion of Finishers		χ^2 Statistic	P-value
Boston	60% and above	$P_M = 32.94\%$	$P_F = 29.76\%$	24.6295	p < 0.0001
	70% and above	$P_M = 5.91\%$	$P_F = 3.19\%$	85.2406	p < 0.0001
	40% and below	$P_M = 6.54\%$	$P_F = 5.18\%$	17.2027	p < 0.0001
Chicago	60% and above	$P_M = 21.74\%$	$P_F = 17.23\%$	122.8513	p < 0.0001
	70% and above	$P_M = 5.50\%$	$P_F = 3.23\%$	114.9892	p < 0.0001
	40% and below	$P_M = 12.08\%$	$P_F = 10.86\%$	13.7648	p = 0.0002
New York	60% and above	$P_M = 21.40\%$	$P_F = 19.91\%$	16.0790	p < 0.0001
	70% and above	$P_M = 4.23\%$	$P_F = 3.41\%$	21.5090	p < 0.0001
	40% and below	$P_M = 7.44\%$	$P_F = 7.84\%$	2.7916	p = 0.0948

Source: (Boston Marathon, 2012; Chicago Marathon, 2013; New York Marathon, 2013)

We also compared the proportion of male and female finishers with age/gender graded scores of seventy percent and above for each race. As with the previous test, we found that the proportion of male finishers was significantly larger than the proportion of female finishers with age/gender graded scores of seventy percent and above. All three races had p-values less than 0.0001, which are highly significant when compared to the significance level of 0.05.

Finally, we compared the proportion of male and female race finishers with age/gender graded scores of forty percent and lower for each race. Unlike the previous tests for proportion, the only races that have a highly significant difference between the male and female proportions are Boston, with a p-value less than 0.00001, and Chicago, with a p-value equal to 0.0002. Both of these races have a larger proportion of male finishers with an age/gender graded score of forty percent or below than the proportion of female race finishers. The result from the New York Marathon was marginally significant, with a p-value of 0.0948, which is less than a significance level of 0.10 but larger than our chosen significance level of 0.05. The New York Marathon is the only result to have the proportion of female runners finishing with an age/gender graded score of forty percent or lower larger than the corresponding proportion of male runners.

Table-3. Descriptive Statistics for Major Marathons

Marathon	Mean		Median		Variance		Std Dev	
Boston	M: 55.6	F: 55.5	M: 56.2	F: 56.3	M: 91.7	F: 73.4	M: 9.57	F: 8.57
Chicago	M: 52.2	F: 51.2	M: 51.7	F: 50.5	M: 108.4	F: 86.4	M: 10.4	F: 9.30
New York	M: 53.0	F: 52.6	M: 52.8	F: 52.1	M: 85.9	F: 83.2	M: 9.27	F: 9.12

Source: (Boston Marathon, 2012; Chicago Marathon, 2013; New York Marathon, 2013)

To further explore the differences between the male and female race finisher age/gender score populations, we calculated the sample means, medians, and variances for the age/gender adjusted scores of males and females for each marathon that requires prequalification to participate. These values are given in Table 3. The F-test Fisher (1924) and Levene (1960) were conducted to test for differences in variances. The t-test, assuming non-equal variances is used to test for differences in means (Weiss, 2012). The Wilcoxon test is used to test for differences in medians and is included here for completeness even though the assumption of equal variances may not be satisfied (Wilcoxon, 1945; Mann and Whitney, 1947). The Kolmogorov-Smirnov (Kolmogorov, 1933; Smirnov, 1948) test is used to test for differences in distributions. The p-value results are given in Table 4. The mean and median age/gender adjusted scores for males was found to be significantly higher than for women for the Chicago and New York Marathons, but not for Boston. The variances of male scores were found to be significantly higher than the variances for the female scores using both the F-test and the Levene test. The distributions of male age/gender adjusted scores were found to be significantly different than the age/gender scores for females in all of the races.

Table-4. Major Marathon Test Results

Marathon	F Test	Levene's Test	T Test	Wilcoxon	KS Test
Boston	p<0.0001	p < 0.0001	p = 0.2948	p = 0.3944	p < 0.0001
Chicago	p<0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
New York	p=0.0142	p = 0.0202	p < 0.0001	p < 0.0001	p < 0.0001

Source: (Boston Marathon, 2012; Chicago Marathon, 2013; New York Marathon, 2013)

For this paper, we wish to combine the test results to see if the overall results follow a similar pattern. To do this, we will be testing the following hypotheses:

H_0: All Separate Null Hypotheses are True

H_a: At Least one of the Separate Alternative Hypotheses is True

To test this hypothesis, we conducted a meta-analysis on the overall results using Stouffer's Z-score method. Stouffer's Z-score method combines p-values of several hypothesis tests (Stouffer *et al.*, 1949). The test statistic is calculated using the standard normal distribution based on the individual p-values and the individual sample sizes for each hypothesis test. The test statistic is defined as

$$Z_i = \phi^{-1}(1 - p_i) \qquad w_i = {}^{n_i}\!/_{\sum_{i=1}^{k} n_i}$$

$$Z = \frac{\sum_{i=1}^{k} w_i Z_i}{\sqrt{\sum_{i=1}^{k} w_i^2}}$$

where ϕ is defined as the standard normal cumulative distribution function, Z_i denotes the Z-score based on the ith p-value, and w_i denotes the weight based on the sample size of the ith hypothesis test. The null hypothesis is

rejected when the test statistic is larger than the critical value from the standard normal distribution based on the chosen level of significance (Stouffer *et al.*, 1949).

We converted the two-sided p-values into one-sided p-values that answer the alternative hypothesis of male results larger than female results; the test for proportions of less than or equal to forty percent was converted to answer the alternative hypothesis of female results larger than male results. These conversions were performed for every statistical test except the test of distributions which has a two-sided alternative hypothesis. After the p-values were converted, we performed a meta-analysis to combine the p-values of the three qualifying marathons using Stouffer's Z-score method. For these combinations, we used the value of 0.99999 for any p-values equal to 1.000 and the value of 0.00009 for any p-values less than 0.0001; these values were chosen for convenience of calculation and to keep the calculation for breaking down as we were using computer programs to calculate the final Z value.

Table-5. Meta-Analysis for Major Marathon Results

	Age Graded Scores	
Variable	**Z value**	**P value**
Proportions		
60%	6.181	p < 0.0001
70%	6.181	p < 0.0001
40%: M < F	-2.276	p = 0.9886
40%: M > F		p = 0.0114
Variances		
F Test	5.212	p < 0.0001
Levene's Test	5.115	p < 0.0001
Means	5.312	p < 0.0001
Medians	4.700	p < 0.0001
Distributions	6.181	p < 0.0001

Source: (Boston Marathon, 2012; Chicago Marathon, 2013; New York Marathon, 2013)

For the proportions of race finishers, at the sixty percent and higher (based on age/gender adjusted scores) and the seventy percent and higher levels, the proportion of male race finishers was significantly higher than the proportion of female race finishers for both of these levels. For the proportion of race finishers at the forty percent and lower level the interpretation is a bit different. We hypothesized that the proportion of female finishers with age/gender adjusted scores lower than 40% was larger than the proportion of male finishers with age/gender adjusted scores lower than 40%. We actually found the opposite to be true. The proportion of male finishers having age/gender adjusted scores lower than 40% was higher than the proportion of female finishers. The variance of age/gender adjusted scores, the mean age/gender adjusted score, and the median age/gender adjusted score for the male race finishers were significantly higher than all of these for women. Finally, the distribution for the male finisher age/gender adjusted scores was significantly different from the distribution for the female finisher age/gender adjusted scores.

4. NON-QUALIFYING RACES – PROPORTION RESULTS

For the larger part of our results, we examined eight races from each of the popular running distances: marathon, half marathon, ten kilometers (10k), and five kilometer (5k). For each race included in our dataset, we ran several statistical tests for the distribution of age/gender adjusted scores as well as the statistical tests for the distribution of ages. We included the age, time, and gender of each race finisher aged fourteen years and older. This information was used to calculate the age/gender adjusted score for each finisher based on the WMA age grading tables (World Master's Athletics, 2010). All races were contested in the year 2013.

We included races of varying sizes from each distance in order to get a more complete picture. Tables 6-9, given below, give the total number of finishers, the total number of male finishers, and the total number of female

finishers for each race at a given distance with Table 6 for marathons, Table 7 for half marathons, Table 8 for 10k races, and Table 9 for 5k races.

Table-6. Total Finishers – Marathon

Marathon	Total	Male	Female
Air Force	3,297	2,232	1,065
Grandma's	5,618	3,275	2,343
Los Angeles	18,691	11,287	7,404
Ogden	2,530	1,322	1,208
Philadelphia	10,913	6,017	4,896
San Francisco	5,823	3,905	1,918
Surf City USA	2,305	1,392	913
Twin Cities	8,852	4,924	3,928

Source: (Airforce Marathon, 2013; Grandma's Marathon, 2013; Los Angeles Marathon, 2013; Philadelphia Marathon, 2013; San Francisco Marathon, 2013; Surf City Marathon and Half-Marathon, 2013; Twin Cities Marathon, 2013; Ogden Marathon, 2015).

The eight marathons that we included in our research are the following: The United States Air Force Marathon on Wright-Patterson AFB, OH, Grandma's Marathon in Duluth, MN, ASICS LA, Zions Bank Ogden, GORE-TEX Philadelphia, Wipro San Francisco, Surf City USA in Huntington Beach, CA, and Medtronic Twin Cities in Minneapolis and St. Paul, MN. For each of these races, the total number of male finishers was greater than the total number of female finishers (Table 6)

Table-7. Total Finishers – Half Marathon

Half Marathon	Total	Male	Female
Berkeley	4,667	2,152	2,515
Brooklyn	21,423	10,822	10,601
Capital City	8,102	3,387	4,715
Columbus	10,200	3,672	6,528
Duluth	6,802	2,850	3,952
Manhattan	4,923	2,981	1,942
New York City	14,531	6,919	7,612
Surf City	14,698	5,881	8,817

Source: (Berkley Half-Marathon, 2013; Brooklyn Half-Marathon, 2013; Capital City Half-Marathon, 2013; Columbus Half-Marathon, 2013; Duluth Half-Marathon, 2013; Manhattan Half-Marathon, 2013; New York City Half-Marathon, 2013; Surf City Marathon and Half-Marathon, 2013).

The half marathons that we included are the following: The Berkeley Half Marathon, New York Road Runners (NYRR) Brooklyn Half Marathon, Capital City Half Marathon in Columbus, OH, Nationwide Children's Hospital Columbus, Garry Bjorklund Half Marathon in Duluth, MN, Manhattan Half-Marathon, New York City (NYC) Half, Surf City USA in Huntington Beach, CA. The total number of female finishers was greater than the number of male finishers for five of the eight races, almost equal in two of the eight races, and the total number of male finishers was greater than the total number of female finishers for one of the races (see Table 7).

The ten kilometer races that we included are the following: The Cooper River Bridge Run in Mount Pleasant, SC, Crescent City Classic in New Orleans, LA, Radio Fargo Moorhead 10k in Fargo, ND, NYRR Joe Kleinerman Classic, Manhattan Beach in Manhattan Beach, CA, NYRR 5-Borough Series: Queens 10k, NYRR Scotland Run 10k in Manhattan, and the UAE Healthy Kidney 10k in Central Park, NYC. The total number of female finishers

was greater than the number of male finishers for three of the eight races, almost equal in two of the races, and the total number of male finishers were greater than the total number of female finishers for three races (see Table 8)

Table-8. Total Finishers – Ten Kilometer

Ten K	Total	Male	Female
Cooper River	30,737	12,625	18,112
Crescent City	17,016	7,909	9,107
Fargo	3,479	958	2,521
Joe Kleinerman	3,137	1,714	1,423
Manhattan Beach	3,455	1,716	1,739
Queens	6,479	3,663	2,816
Scotland Run	7,255	3,936	3,319
UAE Healthy Kidney	5,854	3,156	2,698

Source: (Cooper River 10K, 2013; Crescent City 10K, 2013; Fargo 10K, 2013; Joe Kleinerman 10K, 2013; Manhattan Beach 10K, 2013; Queens 10K, 2013; Scotland Run 10K, 2013; UAE Healthy Kidney 10K, 2013).

Table-9. Total Finishers – Five Kilometer

Five K	Total	Male	Female
Boston	5,477	2,422	3,055
Chelsea's Run	1,626	646	980
Gasparilla	9,915	3,961	5,954
McGuire's	11,977	5,237	6,740
MLB All Star	4,722	2,469	2,253
NYRR Dash	7,993	3,598	4,395
OC	1,881	644	1,237
Percy Sutton	3,464	1,711	1,753

Source: (Boston 5K, 2013; Chelsea's Run 5K, 2013; Gasparilla 5K, 2013; McGuire's 5K, 2013; MLB All Star 5K, 2013; NYRR Dash 5K, 2013; OC 5K, 2013; Percy Sutton 5K, 2013).

The five kilometer races that we included are the following: the Boston Athletic Association 5k in Boston, Publix Super Markets Gasparilla Distance Classic in Tampa Bay, FL, MLB All-Star Run Benefitting Sandy Relief in Prospect Park, Brooklyn, NY, OC Marathon 5k in Newport Beach, CA, Finish Chelsea's Run in San Diego, CA, McGuire's St. Patrick's Day 5k in Pensacola, FL, NYRR Dash to the Finish Line 5k in NYC, and the Percy Sutton Harlem 5k Run. The total number of female finishers is greater than the number of male finishers for six of the eight races, and almost equal in the other two races (see Table 9).

For the eight marathon races we examined in this paper, we looked at the distributions of age/gender graded scores for male and female race finishers and the distributions of the ages for male and female race finishers. We first considered the age/gender graded scores of each of the marathons individually and tested whether or not there was a difference in the proportions of males and females finishing with an age/gender graded score of 60% and above using a chi-square test. Significant differences were found in 6 or the 8 marathons (α=0.05) with the proportion of male finishers with an age/gender graded score of 60% of above being significantly greater than female finishers in 5 of the marathons and the proportion of female finishers with an age/gender graded score being significantly greater in one of the marathons (Ogden). Results are given in Table 10.

Table-10. Proportion of Marathon Finishers – 60% and Above

Marathon	Proportion of Finishers		χ² Statistic	P-value
Air Force	$P_M = 15.82\%$	$P_F = 14.84\%$	0.5284	p = 0.4673
Duluth	$P_M = 34.47\%$	$P_F = 29.58\%$	14.9397	p = 0.0001
Los Angeles	$P_M = 10.22\%$	$P_F = 8.04\%$	25.2348	p < 0.0001
Ogden	$P_M = 22.24\%$	$P_F = 25.75\%$	4.2644	p = 0.0389
Philadelphia	$P_M = 27.57\%$	$P_F = 24.14\%$	16.4904	p < 0.0001
San Francisco	$P_M = 15.75\%$	$P_F = 13.35\%$	5.8330	p = 0.0157
Surf City	$P_M = 20.91\%$	$P_F = 21.25\%$	0.0391	p = 0.8432
Twin Cities	$P_M = 26.20\%$	$P_F = 23.47\%$	8.6604	p = 0.0033

Source: (Airforce Marathon, 2013; Grandma's Marathon, 2013; Los Angeles Marathon, 2013; Philadelphia Marathon, 2013; San Francisco Marathon, 2013; Surf City Marathon and Half-Marathon, 2013; Twin Cities Marathon, 2013; Ogden Marathon, 2015).

Tests were also conducted between the proportion of male finishers with age/gender graded scores at or above 70% compared to the proportion of female finishers. Significant differences were found in 4 of the 8 marathons with the proportion of male finishers higher in all of the four.

We next tested to see whether or not the proportion of male finishers with an age/gender graded score of 40% or lower was different than this proportion for female finishers. Results are given in Table 11.

The proportion of males was significantly lower in Los Angeles and marginally significantly lower in Duluth with marginal significance meaning the p-value was between 0.05 and 0.10. The proportion of females was significantly lower in Philadelphia, San Francisco, and the Twin Cities, and marginally significantly lower in Ogden.

Table-11. Proportion of Marathon Finishers – 40% and Below

Marathon	Proportion of Finishers		χ² Statistic	P-value
Air Force	$P_M = 17.70\%$	$P_F = 18.69\%$	0.4768	p = 0.4899
Duluth	$P_M = 3.93\%$	$P_F = 4.85\%$	2.7581	p = 0.0968
Los Angeles	$P_M = 27.91\%$	$P_F = 29.81\%$	7.8941	p = 0.0050
Ogden	$P_M = 6.51\%$	$P_F = 4.88\%$	3.0706	p = 0.0797
Philadelphia	$P_M = 7.40\%$	$P_F = 5.94\%$	9.0501	p = 0.0026
San Francisco	$P_M = 12.88\%$	$P_F = 10.22\%$	8.6284	p = 0.0033
Surf City	$P_M = 10.42\%$	$P_F = 10.51\%$	0.0057	p = 0.9400
Twin Cities	$P_M = 5.54\%$	$P_F = 3.62\%$	18.1991	p < 0.0001

Source: (Airforce Marathon, 2013; Grandma's Marathon, 2013; Los Angeles Marathon, 2013; Philadelphia Marathon, 2013; San Francisco Marathon, 2013; Surf City Marathon and Half-Marathon, 2013; Twin Cities Marathon, 2013; Ogden Marathon, 2015)

4.1. Half-Marathon Results-Proportions

Tests were conducted comparing the proportions of males and females with age/gender graded scores of 60% or higher for each of the 8 half-marathons in the sample. The proportion of male finishers with an age/gender graded score of 60% or higher was significantly higher than the proportion of female finishers with an age-graded score of 60% or higher for all of the half-marathons in the sample. The results are given in Table 12.

Table-12. Proportion of Half Marathon Finishers – 60% and Above

Half Marathon	Proportion of Finishers		χ² Statistic	P-value
Berkeley	$P_M = 24.54\%$	$P_F = 18.29\%$	27.1022	p < 0.0001
Brooklyn	$P_M = 27.00\%$	$P_F = 19.00\%$	193.3983	p < 0.0001
Capital City	$P_M = 16.18\%$	$P_F = 10.80\%$	50.3700	p < 0.0001
Columbus	$P_M = 17.10\%$	$P_F = 11.53\%$	62.2280	p < 0.0001
Duluth	$P_M = 27.72\%$	$P_F = 19.66\%$	60.6467	p < 0.0001
Manhattan	$P_M = 28.92\%$	$P_F = 23.22\%$	19.4872	p < 0.0001
New York City	$P_M = 28.11\%$	$P_F = 21.69\%$	80.2565	p < 0.0001
Surf City	$P_M = 13.98\%$	$P_F = 9.62\%$	8.6604	p < 0.0001

Source: (Berkley Half-Marathon, 2013; Brooklyn Half-Marathon, 2013; Capital City Half-Marathon, 2013; Columbus Half-Marathon, 2013; Duluth Half-Marathon, 2013; Manhattan Half-Marathon, 2013; New York City Half-Marathon, 2013; Surf City Marathon and Half-Marathon, 2013)

Tests were also conducted comparing the proportion of male finishers with an age/gender graded score of 70% of higher to the proportion of female finishers with an age-graded score of 70% of higher. The proportion of males was significantly higher in all of the 8 races.

Table-13. Proportion of Half Marathon Finishers – 40% and Below

Half Marathon	Proportion of Finishers		χ^2 Statistic	P-value
Berkeley	$P_M = 9.25\%$	$P_F = 11.25\%$	5.0353	p = 0.0248
Brooklyn	$P_M = 3.66\%$	$P_F = 4.00\%$	1.6858	p = 0.1942
Capital City	$P_M = 8.77\%$	$P_F = 12.45\%$	27.4730	p < 0.0001
Columbus	$P_M = 11.63\%$	$P_F = 14.29\%$	14.4303	p = 0.0001
Duluth	$P_M = 5.44\%$	$P_F = 6.48\%$	3.1494	p = 0.0760
Manhattan	$P_M = 3.69\%$	$P_F = 3.50\%$	0.1199	p = 0.7292
New York City	$P_M = 3.93\%$	$P_F = 4.16\%$	0.5071	p = 0.4764
Surf City	$P_M = 20.47\%$	$P_F = 25.03\%$	41.1349	p < 0.0001

Source: (Berkley Half-Marathon, 2013; Brooklyn Half-Marathon, 2013; Capital City Half-Marathon, 2013; Columbus Half-Marathon, 2013; Duluth Half-Marathon, 2013; Manhattan Half-Marathon, 2013; New York City Half-Marathon, 2013; Surf City Marathon and Half-Marathon, 2013)

Tests were conducted comparing the proportion of male finishers who had age/gender adjusted scores at or up 40% compared to the proportion of female finishers who had age/gender adjusted scores at or up 40% for each of the half-marathons. The proportion of male finishers was significantly lower in the Berkeley, Capital City, Columbus, and Surf City half-marathons and marginally significantly lower in the Duluth half-marathon. Results are given in Table 13.

4.2. Ten K Results- Proportions

Tests were conducted comparing the proportion of male finishers with an age/gender graded score at or above 60% with the proportion of female finishers with an age/gender graded score at or above 60%. The proportion of male finishers with these scores was significantly higher for each of the 8 races. The results are given in Table 14. When comparing male and female finishers with an age/gender graded score of 70% of higher, the proportion of male finishers was always significantly larger for each of the races.

Table-14. Proportion of Ten K Finishers – 60% and Above

Ten K Race	Proportion of Finishers		χ^2 Statistic	P-value
Cooper River	$P_M = 10.50\%$	$P_F = 5.20\%$	306.0013	p < 0.0001
Crescent City	$P_M = 11.42\%$	$P_F = 4.25\%$	310.4056	p < 0.0001
Fargo	$P_M = 10.75\%$	$P_F = 5.43\%$	30.5590	p < 0.0001
Joe K	$P_M = 31.45\%$	$P_F = 23.47\%$	24.6230	p < 0.0001
M Beach	$P_M = 24.88\%$	$P_F = 17.42\%$	28.8411	p < 0.0001
Queens	$P_M = 21.02\%$	$P_F = 12.04\%$	90.5488	p < 0.0001
Scotland Run	$P_M = 31.28\%$	$P_F = 19.67\%$	126.0492	p < 0.0001
UAE	$P_M = 23.92\%$	$P_F = 12.97\%$	113.8950	p < 0.0001

Source: (Cooper River 10K, 2013; Crescent City 10K, 2013; Fargo 10K, 2013; Joe Kleinerman 10K, 2013; Manhattan Beach 10K, 2013; Queens 10K, 2013; Scotland Run 10K, 2013; UAE Healthy Kidney 10K, 2013)

Tests were conducted comparing the proportions of male finishers at or below 40% with the proportion of female finishers at or below 40% for each of the 10K races. The proportion of male finishers was found to be significantly lower in 5 of the 8 races. Results are given in Table 15.

Table-15. Proportion of Ten K Finishers – 40% and Below

Ten K Race	Proportion of Finishers		χ² Statistic	P-value
Cooper River	P_M= 30.92%	P_F = 42.35%	413.6307	p < 0.0001
Crescent City	P_M= 36.54%	P_F = 52.86%	455.0133	p < 0.0001
Fargo	P_M= 19.83%	P_F = 24.87%	9.8061	p = 0.0017
Joe K	P_M = 3.38%	P_F = 3.65%	0.1680	p = 0.6819
M Beach	P_M= 10.31%	P_F = 14.55%	14.2087	p = 0.0002
Queens	P_M = 5.65%	P_F = 6.50%	2.0212	p = 0.1551
Scotland Run	P_M = 2.01%	P_F = 4.37%	33.5640	p < 0.0001
UAE	P_M = 5.07%	P_F = 5.74%	1.3028	p = 0.2537

Source: (Cooper River 10K, 2013; Crescent City 10K, 2013; Fargo 10K, 2013; Joe Kleinerman 10K, 2013; Manhattan Beach 10K, 2013; Queens 10K, 2013; Scotland Run 10K, 2013; UAE Healthy Kidney 10K, 2013)

4.3. Five K Results- Proportions

Tests were conducted comparing the proportion of male finishers with an age/gender graded score of 60% or higher with the proportion of female finishers with an age/gender graded score of 60% or higher for each of the 8 5k races in the sample. The proportion of male finishers was found to be significantly higher for all of the races. Results are given in Table 16. Similar results were found when comparing the proportions finishing at 70% or higher.

Table-16. Proportion of Five K Finishers – 60% and Above

Five K Race	Proportion of Finishers		χ² Statistic	P-value
Boston	P_M= 31.30%	P_F = 21.51%	67.5867	p < 0.0001
Chelsea's Run	P_M= 19.50%	P_F = 11.43%	20.3242	p < 0.0001
Gasparilla	P_M = 5.78%	P_F = 2.99%	47.0962	p < 0.0001
McGuire's	P_M = 4.45%	P_F = 2.58%	31.3104	p < 0.0001
MLB All Star	P_M= 14.14%	P_F = 10.12%	17.7082	p < 0.0001
NYRR Dash	P_M= 14.73%	P_F = 10.24%	37.1007	p < 0.0001
OC	P_M = 9.63%	P_F = 4.53%	18.7367	p < 0.0001
Percy Sutton	P_M= 28.93%	P_F = 17.17%	67.6577	p < 0.0001

Source: (Boston 5K, 2013; Chelsea's Run 5K, 2013; Gasparilla 5K, 2013; McGuire's 5K, 2013; MLB All Star 5K, 2013; NYRR Dash 5K, 2013; OC 5K, 2013; Percy Sutton 5K, 2013)

Tests were also conducted comparing the proportion of male finishers with an age/gender graded score of 40% or lower to the proportion of female finishers with an age-graded score of 40% or lower. The proportion of males was significantly lower in 6 of the races and marginally significantly lower in one of the races. Results are given in Table 17.

Table-17. Five K Finishers – 40% and Below

Five K Race	Proportion of Finishers		χ² Statistic	P-value
Boston	P_M = 8.55%	P_F = 8.12%	0.3261	p = 0.5680
Chelsea's Run	P_M= 23.53%	P_F = 29.39%	6.7700	p = 0.0093
Gasparilla	P_M= 44.23%	P_F = 53.43%	80.4445	p < 0.0001
McGuire's	P_M= 51.80%	P_F = 61.91%	123.2474	p < 0.0001
MLB All Star	P_M= 15.92%	P_F = 17.75%	2.8441	p = 0.0917
NYRR Dash	P_M= 16.12%	P_F = 14.04%	6.7311	p = 0.0095
OC	P_M= 43.01%	P_F = 50.69%	9.9933	p = 0.0016
Percy Sutton	P_M = 7.77%	P_F = 12.21%	18.8893	p < 0.0001

Source: (Boston 5K, 2013; Chelsea's Run 5K, 2013; Gasparilla 5K, 2013; McGuire's 5K, 2013; MLB All Star 5K, 2013; NYRR Dash 5K, 2013; OC 5K, 2013; Percy Sutton 5K, 2013)

5. NON-QUALIFYING RACES- TESTS FOR DIFFERENCES IN AGE/GENDER GRADED SCORES AND AGES – MEANS, MEDIANS, VARIANCES

To further determine where the differences lie between the distributions of male and female finishers we ran several tests. We tested the equality of the means, medians, variances, and overall distributions of age/gender graded scores for both male and female populations, and then ages for both populations. We used the two-sample T test to test for differences in means (Weiss, 2012) the F test to test for differences in variances (Fisher, 1924) Levene's test to test for differences in variances (Levene, 1960) the Wilcoxon test to test for differences in medians (Wilcoxon, 1945; Mann and Whitney, 1947) and the KS test to test for differences in distributions (Kolmogorov, 1933; Smirnov, 1948). The F test assumes underlying normal populations, whereas, Levene's test is more robust to this assumption. We will use the results of these tests to draw our conclusions about the differences between the distributions of male and female distance runners.

5.1. Marathon Results

The sample means, medians, variances, and standard deviations of men's and women's age/gender adjusted scores for each of the non-qualifying marathons in our sample are given in Table 18.

Table-18. Descriptive Statistics for Age/Gender Graded Score – Marathon

Marathon	Mean		Median		Variance		Standard Dev	
Air Force	M: 49.6	F: 49.3	M: 49.0	F: 48.5	M: 96.8	F: 95.1	M: 9.84	F: 9.75
Duluth	M: 56.4	F: 55.3	M: 55.8	F: 54.6	M: 113.2	F: 89.3	M: 10.64	F: 9.45
Los Angeles	M: 46.6	F: 45.8	M: 45.7	F: 44.7	M: 102.5	F: 90.7	M: 10.12	F: 9.53
Ogden	M: 53.2	F: 54.2	M: 52.9	F: 54.1	M: 83.2	F: 74.9	M: 9.12	F: 8.66
Philadelphia	M: 54.1	F: 53.7	M: 53.4	F: 53.1	M: 102.1	F: 83.3	M: 10.10	F: 9.13
San Francisco	M: 50.6	F: 50.4	M: 49.9	F: 49.6	M: 84.3	F: 73.4	M: 9.18	F: 8.57
Surf City	M: 52.0	F: 52.1	M: 51.8	F: 51.4	M: 91.9	F: 93.6	M: 9.58	F: 9.68
Twin Cities	M: 54.4	F: 54.1	M: 53.4	F: 53.4	M: 97.6	F: 81.0	M: 9.88	F: 9.00

Source: (Airforce Marathon, 2013; Grandma's Marathon, 2013; Los Angeles Marathon, 2013; Philadelphia Marathon, 2013; San Francisco Marathon, 2013; Surf City Marathon and Half-Marathon, 2013; Twin Cities Marathon, 2013; Ogden Marathon, 2015)

Table-19. Test Results for Age/Gender Graded Score – Marathon

Marathon	F Test	Levene's Test	T Test	Wilcoxon	KS Test
Air Force	p = 0.7405	p = 0.7131	p = 0.4022	p = 0.3675	p = 0.8068
Duluth	p < 0.0001	p < 0.0001	p = 0.0001	p = 0.0005	p = 0.0002
Los Angeles	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Ogden	p = 0.0634	p = 0.0503	p = 0.0033	p = 0.0018	p = 0.0054
Philadelphia	p < 0.0001	p < 0.0001	p = 0.0258	p = 0.1720	p = 0.0002
San Francisco	p = 0.0005	p = 0.0006	p = 0.4024	p = 0.3143	p = 0.0539
Surf City	p = 0.7469	p = 0.7330	p = 0.8384	p = 0.9995	p = 0.9649
Twin Cities	p < 0.0001	p < 0.0001	p = 0.0905	p = 0.3489	p = 0.0096

Source: (Airforce Marathon, 2013; Grandma's Marathon, 2013; Los Angeles Marathon, 2013; Philadelphia Marathon, 2013; San Francisco Marathon, 2013; Surf City Marathon and Half-Marathon, 2013; Twin Cities Marathon, 2013; Ogden Marathon, 2015)

We considered two tests for the differences in the variances between the male and female finishers, the F test and Levene's test. It was found that the variances in male age/gender adjusted scores were significantly larger in 5 of the 8 marathons than the variances in female age/gender adjusted running times. For the Ogden marathon, the variance of the male age/gender adjusted scores was found to be marginally significantly higher. In two of the marathons, Air Force, and Surf City, the variances were not found to be significantly different. The t-test, assuming non-equal variances, was used to test for differences in mean age/gender adjusted scores between males and females for each of the marathons. Significant differences were found in 4 of the 8 marathons with the means

for males significantly larger than the means for females in three of these. A marginally significant difference was found in the mean age/gender adjusted scores between males and females for the Twin Cities marathon with the mean for males larger. The p-values for the Wilcoxon test for differences in medians are given in the table also for comparison purposes. However, it should be noted that the Wilcoxon test does assume equality of variances which may not be the case.

Table-20. Descriptive Statistics for Age – Marathon

Marathon	Mean		Median		Variance		Standard Dev	
Air Force	M: 40.0	F: 37.7	M: 39.0	F: 37.0	M: 136.9	F: 116.4	M: 11.70	F: 10.79
Duluth	M: 39.3	F: 35.1	M: 38.0	F: 34.0	M: 155.7	F: 110.5	M: 12.48	F: 10.51
Los Angeles	M: 36.7	F: 34.1	M: 36.0	F: 33.0	M: 182.3	F: 153.6	M: 13.50	F: 12.39
Ogden	M: 39.9	F: 37.4	M: 39.0	F: 37.0	M: 105.0	F: 82.2	M: 10.25	F: 9.07
Philadelphia	M: 38.3	F: 34.8	M: 38.0	F: 33.0	M: 120.4	F: 91.8	M: 10.97	F: 9.58
San Francisco	M: 37.7	F: 35.1	M: 36.0	F: 33.0	M: 113.8	F: 95.4	M: 10.67	F: 9.77
Surf City	M: 42.6	F: 38.7	M: 42.0	F: 39.0	M: 134.4	F: 105.7	M: 11.59	F: 10.28
Twin Cities	M: 39.0	F: 34.7	M: 38.0	F: 33.0	M: 130.9	F: 91.6	M: 11.44	F: 9.57

Source: (Airforce Marathon, 2013; Grandma's Marathon, 2013; Los Angeles Marathon, 2013; Philadelphia Marathon, 2013; San Francisco Marathon, 2013; Surf City Marathon and Half-Marathon, 2013; Twin Cities Marathon, 2013; Ogden Marathon, 2015)

The results for the Wilcoxon test generally agree with the results from the t-test except for the Philadelphia and Twin Cities marathons in which no significant differences are found. To test the equality of the distributions of the age/gender graded scores for the male and female finishers, we performed a KS test and found that six of the eight races had statistically significant distributions between the two genders. The Air Force and the Surf City USA marathons had highly non-significant differences.

The descriptive statistics for the ages of participants in each of the marathons is given in Table 20. Test results for ages are given in Table 21. The mean age of the male finishers is significantly larger than the mean age of the female finishers in all eight races. The median age of the race finishers and the variance of the ages for the race finishers are also both significantly larger for males in all of the eight marathons considered.

Table-21. Test Results for Age – Marathon

Marathon	F Test	Levene's Test	T Test	Wilcoxon	KS Test
Air Force	p = 0.0023	p = 0.0005	p < 0.0001	p < 0.0001	p < 0.0001
Duluth	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Los Angeles	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Ogden	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Philadelphia	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
San Francisco	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Surf City	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Twin Cities	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001

Source: (Airforce Marathon, 2013; Grandma's Marathon, 2013; Los Angeles Marathon, 2013; Philadelphia Marathon, 2013; San Francisco Marathon, 2013; Surf City Marathon and Half-Marathon, 2013; Twin Cities Marathon, 2013; Ogden Marathon, 2015)

5.2. Half-Marathon Results

The descriptive statistics for the age/gender graded scores for each of the 8 half-marathons are given in Table 22. Test results for age/gender graded scores are given in Table 23.

Table-22. Descriptive Statistics for Age/Gender Graded Score – Half Marathon

Half Marathon	Mean		Median		Variance		Standard Dev	
Berkeley	M: 53.5	F: 52.0	M: 53.0	F: 52.0	M: 104.8	F: 88.3	M: 10.24	F: 9.40
Brooklyn	M: 55.1	F: 53.6	M: 54.7	F: 53.3	M: 76.3	F: 61.9	M: 8.74	F: 7.87
Capital City	M: 51.8	F: 49.7	M: 51.7	F: 49.7	M: 76.9	F: 70.4	M: 8.77	F: 8.39
Columbus	M: 51.2	F: 49.5	M: 51.0	F: 49.5	M: 98.1	F: 78.8	M: 9.90	F: 8.88
Duluth	M: 55.6	F: 52.8	M: 54.0	F: 51.9	M: 149.5	F: 91.2	M:12.23	F: 9.55
Manhattan	M: 55.2	F: 54.1	M: 55.1	F: 53.7	M: 76.0	F: 68.1	M: 8.72	F: 8.25
New York City	M: 55.3	F: 54.0	M: 54.5	F: 53.7	M: 89.6	F: 74.4	M: 9.47	F: 8.63
Surf City	M: 48.8	F: 47.1	M: 48.4	F: 46.6	M: 109.4	F: 93.8	M: 10.46	F: 9.68

Source: (Berkley Half-Marathon, 2013; Brooklyn Half-Marathon, 2013; Capital City Half-Marathon, 2013; Columbus Half-Marathon, 2013; Duluth Half-Marathon, 2013; Manhattan Half-Marathon, 2013; New York City Half-Marathon, 2013; Surf City Marathon and Half-Marathon, 2013)

The Brooklyn half marathon and the Manhattan half marathon have a larger number of total male finishers than total female finishers but the six other races that were examined have a larger number of total female finishers than total male finishers. For all eight races, the mean age/gender adjusted score for male finishers is significantly higher than the mean age/gender adjusted score for female finishers. This trend holds true for the median age/gender adjusted scores and the variance for the age/gender adjusted scores between the male and female half marathon finishers.

Table-23. Test Results for Age/Gender Graded Score – Half Marathon

Half Marathon	F Test	Levene's Test	T Test	Wilcoxon	KS Test
Berkeley	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p = 0.0006
Brooklyn	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Capital City	p = 0.0055	p = 0.0067	p < 0.0001	p < 0.0001	p < 0.0001
Columbus	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Duluth	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Manhattan	p = 0.0076	p = 0.0074	p < 0.0001	p < 0.0001	p < 0.0001
New York City	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Surf City	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001

Source: (Berkley Half-Marathon, 2013; Brooklyn Half-Marathon, 2013; Capital City Half-Marathon, 2013; Columbus Half-Marathon, 2013; Duluth Half-Marathon, 2013; Manhattan Half-Marathon, 2013; New York City Half-Marathon, 2013; Surf City Marathon and Half-Marathon, 2013)

Descriptive statistics for ages for each of the half-marathons in the samples are given in Table 24. The test results for ages for half-marathons are given in Table 25.

Table-24. Descriptive Statistics for Age – Half Marathon

Half Marathon	Mean		Median		Variance		Standard Dev	
Berkeley	M: 36.8	F: 35.6	M: 35.0	F: 34.0	M: 110.7	F: 102.6	M:10.52	F:10.13
Brooklyn	M: 36.3	F: 33.3	M: 34.0	F: 31.0	M: 89.2	F: 66.2	M: 9.44	F: 8.14
Capital City	M: 38.2	F: 34.5	M: 37.0	F: 33.0	M: 123.1	F: 106.6	M:11.10	F:10.33
Columbus	M: 37.8	F: 34.3	M: 36.0	F: 33.0	M: 152.2	F: 118.2	M:12.34	F:10.87
Duluth	M: 37.3	F: 34.1	M: 35.0	F: 33.0	M: 155.9	F: 118.3	M:12.48	F:10.88
Manhattan	M: 39.8	F: 35.5	M: 38.0	F: 33.5	M: 114.4	F: 94.3	M:10.70	F: 9.71
New York City	M: 38.9	F: 35.0	M: 37.0	F: 33.0	M: 105.3	F: 83.8	M:10.26	F: 9.16
Surf City	M: 40.7	F: 38.6	M: 40.0	F: 38.0	M: 135.2	F: 115.9	M:11.63	F:10.77

Source: (Berkley Half-Marathon, 2013; Brooklyn Half-Marathon, 2013; Capital City Half-Marathon, 2013; Columbus Half-Marathon, 2013; Duluth Half-Marathon, 2013; Manhattan Half-Marathon, 2013; New York City Half-Marathon, 2013; Surf City Marathon and Half-Marathon, 2013)

Table-25. Test Results for Age – Half Marathon

Half Marathon	F Test	Levene's Test	T Test	Wilcoxon	KS Test
Berkeley	p = 0.0656	p = 0.0695	p < 0.0001	p < 0.0001	p < 0.0001
Brooklyn	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Capital City	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Columbus	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Duluth	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Manhattan	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
New York City	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Surf City	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001

Source: (Berkley Half-Marathon, 2013; Brooklyn Half-Marathon, 2013; Capital City Half-Marathon, 2013; Columbus Half-Marathon, 2013; Duluth Half-Marathon, 2013; Manhattan Half-Marathon, 2013; New York City Half-Marathon, 2013; Surf City Marathon and Half-Marathon, 2013)

The Brooklyn half marathon and the Manhattan half marathon have a larger number of total male finishers than total female finishers but the six other races that were examined have a larger number of total female finishers than total male finishers. For all eight races, the mean age of the male finishers is significantly higher than the mean age of the female finishers. This trend holds true for the median age and the variance for the ages of the male and female half marathon finishers with the median age being significantly larger for males in all 8 races and the variance being significantly larger for males in 7 of the 8 races and marginally significantly larger in the Berkeley half-marathon.

5.3. K Results

The descriptive statistics for age/gender adjusted scores based on the samples for each of the 8 ten kilometer races in the sample are given in Table 26. The test results for the age/gender adjusted scores are given in Table 27.

Table-26. Descriptive Statistics for Age/Gender Graded Score – Ten Kilometer

Ten K Race	Mean		Median		Variance		Standard Dev	
Cooper River	M: 45.6	F: 42.4	M: 46.0	F: 42.5	M: 142.2	F: 121.4	M: 11.92	F: 11.02
Crescent City	M: 44.4	F: 39.8	M: 44.7	F: 39.0	M: 163.6	F: 123.0	M: 12.79	F: 11.09
Fargo	M: 48.2	F: 46.2	M: 48.3	F: 46.3	M: 90.7	F: 77.0	M: 9.53	F: 8.77
Joe K	M: 55.9	F: 53.9	M: 55.7	F: 53.3	M: 85.7	F: 78.2	M: 9.26	F: 8.84
M Beach	M: 53.3	F: 50.9	M: 53.1	F: 51.1	M: 108.8	F: 100.3	M: 10.43	F: 10.02
Queens	M: 53.1	F: 50.9	M: 52.5	F: 50.5	M: 75.3	F: 57.1	M: 8.68	F: 7.56
Scotland Run	M: 56.3	F: 54.0	M: 55.4	F: 53.0	M: 89.9	F: 71.3	M: 9.48	F: 8.44
UAE	M: 54.2	F: 51.8	M: 53.7	F: 51.6	M: 89.7	F: 61.7	M: 9.47	F: 7.86

Source: (Cooper River 10K, 2013; Crescent City 10K, 2013; Fargo 10K, 2013; Joe Kleinerman 10K, 2013; Manhattan Beach 10K, 2013; Queens 10K, 2013; Scotland Run 10K, 2013; UAE Healthy Kidney 10K, 2013)

The total number of male finishers was larger than the total number of female finishers for half of the ten kilometer races that we investigated and the total number of female finishers was larger than the total number of male finishers for the other four races. The mean age/gender adjusted score for the male ten k finishers was significantly larger than the mean age/gender adjusted score for the female finishers. This was also true of the median age/gender scores. The differences between the variances for the age/gender adjusted scores for the male and female ten k finishers are significant for six of the eight races that were examined. For the Manhattan Beach ten kilometer race and the Joe Kleinerman ten kilometer race, the results of the F test and the Levene's test are marginally significantly with the variance for males higher. The distributions for age/gender adjusted scores of males and females are all significantly different.

Table-27. Test Results for Age/Gender Graded Score – Ten Kilometer

Ten K Race	F Test	Levene's Test	T Test	Wilcoxon	KS Test
Cooper River	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Crescent City	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Fargo	p = 0.0019	p = 0.0026	p < 0.0001	p < 0.0001	p < 0.0001
Joe K	p = 0.0698	p = 0.0827	p < 0.0001	p < 0.0001	p < 0.0001
M Beach	p = 0.0926	p = 0.0929	p < 0.0001	p < 0.0001	p < 0.0001
Queens	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Scotland Run	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
UAE	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001

Source: (Cooper River 10K, 2013; Crescent City 10K, 2013; Fargo 10K, 2013; Joe Kleinerman 10K, 2013; Manhattan Beach 10K, 2013; Queens 10K, 2013; Scotland Run 10K, 2013; UAE Healthy Kidney 10K, 2013)

The descriptive statistics for ages based on the samples for each of the 8 ten kilometer races are given in Table 28. Test results for ages are given in Table 29 for the ten kilometer races.

Table-28. Descriptive Statistics for Age – Ten Kilometer

Ten K Race	Mean		Median		Variance		Standard Dev	
Cooper River	M: 39.5	F: 37.1	M: 38.0	F: 35.0	M: 184.9	F: 155.9	M: 13.60	F: 12.48
Crescent City	M: 38.6	F: 36.6	M: 37.0	F: 35.0	M: 186.5	F: 156.2	M: 13.66	F: 12.50
Fargo	M: 35.6	F: 36.3	M: 34.0	F: 34.0	M: 178.6	F: 132.7	M: 13.37	F: 11.52
Joe K	M: 40.6	F: 36.2	M: 39.0	F: 34.0	M: 121.7	F: 110.8	M: 11.03	F: 10.53
M Beach	M: 42.3	F: 38.8	M: 42.0	F: 38.0	M: 213.9	F: 175.8	M: 14.63	F: 13.26
Queens	M: 37.9	F: 34.7	M: 36.0	F: 33.0	M: 98.5	F: 79.2	M: 9.93	F: 8.90
Scotland Run	M: 38.2	F: 34.3	M: 36.0	F: 32.0	M: 108.0	F: 86.8	M: 10.39	F: 9.32
UAE	M: 38.2	F: 34.7	M: 36.0	F: 32.0	M: 109.5	F: 87.5	M: 10.47	F: 9.36

Source: (Cooper River 10K, 2013; Crescent City 10K, 2013; Fargo 10K, 2013; Joe Kleinerman 10K, 2013; Manhattan Beach 10K, 2013; Queens 10K, 2013; Scotland Run 10K, 2013; UAE Healthy Kidney 10K, 2013)

For the Fargo 10k, the mean age for the female finishers is larger than the mean age for the male finishers, and the median age for the female finishers is equal to the median age for the male finishers. For the remaining seven ten kilometer races, the mean and median ages for the male finishers are larger than the mean and median ages for the female finishers. All eight races have the variance for the ages of the male finishers larger than the variance for the ages of the female finishers. The variance is significantly larger for males in 7 of the races and marginally significantly larger in one of the races. The age distributions are all significantly different.

Table-29. Test Results for Age – Ten Kilometer

Ten K Race	F Test	Levene's Test	T Test	Wilcoxon	KS Test
Cooper River	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Crescent City	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Fargo	p < 0.0001	p < 0.0001	p = 0.1975	p >0.05	p = 0.0163
Joe K	p = 0.0642	p = 0.0917	p < 0.0001	p < 0.0001	p < 0.0001
M Beach	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Queens	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Scotland Run	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
UAE	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001

Source: (Cooper River 10K, 2013; Crescent City 10K, 2013; Fargo 10K, 2013; Joe Kleinerman 10K, 2013; Manhattan Beach 10K, 2013; Queens 10K, 2013; Scotland Run 10K, 2013; UAE Healthy Kidney 10K, 2013)

5.4 K Results

Descriptive statistics for the age/gender adjusted scores based on the 8 five kilometer races in the sample are given in Table 30. Test results for age/gender graded scores are given in Table 31.

The total number of male finishers was larger than the total number of female finishers for the McGuire's St Patrick's Day five kilometer race but for all other races that were examined the total number of female finishers was larger than the total number of male finishers. The mean age/gender adjusted score for the male finishers was significantly larger than the mean age/gender adjusted score for the female finishers for 7 of the 8 races. This is also true of the median age/gender adjusted scores. The variance of the age/gender adjusted scores for males was significantly larger than the variance of the age/gender adjusted scores for women in all of the eight races. The distributions of age/gender adjusted scores for males and females were all significantly different.

Table-30. Descriptive Statistics for Age/Gender Graded Score – Five Kilometer

Five K Race	Mean		Median		Variance		Standard Dev	
Boston	M: 54.8	F: 53.0	M: 54.1	F: 52.4	M: 130.7	F: 97.8	M: 11.43	F: 9.89
Chelsea's Run	M: 49.0	F: 46.4	M: 49.6	F: 46.6	M: 156.0	F: 128.2	M: 12.49	F: 11.32
Gasparilla	M: 41.9	F: 39.4	M: 41.7	F: 39.0	M: 129.1	F: 101.5	M: 11.36	F: 10.08
McGuire's	M: 39.6	F: 37.6	M: 39.5	F: 36.3	M: 134.1	F: 105.9	M: 11.58	F: 10.29
MLB All Star	M: 49.3	F: 48.4	M: 49.2	F: 48.6	M: 99.1	F: 84.4	M: 9.95	F: 9.19
NYRR Dash	M: 49.7	F: 49.4	M: 48.6	F: 49.1	M: 110.8	F: 86.5	M: 10.53	F: 9.30
OC	M: 42.4	F: 40.9	M: 42.1	F: 39.9	M: 163.4	F: 112.0	M: 12.78	F: 10.59
Percy Sutton	M: 54.5	F: 51.2	M: 54.4	F: 50.8	M: 107.0	F: 95.8	M: 10.35	F: 9.79

Source: (Boston 5K, 2013; Chelsea's Run 5K, 2013; Gasparilla 5K, 2013; McGuire's 5K, 2013; MLB All Star 5K, 2013; NYRR Dash 5K, 2013; OC 5K, 2013; Percy Sutton 5K, 2013)

Table-31. Test Results for Age/Gender Graded Score – Five Kilometer

Five K Race	F Test	Levene's Test	T Test	Wilcoxon	KS Test
Boston	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Chelsea's Run	p = 0.0057	p = 0.0028	p < 0.0001	p < 0.0001	p < 0.0001
Gasparilla	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
McGuire's	p = 0.0001	p = 0.0003	p = 0.0008	p = 0.0025	p = 0.0008
MLB All Star	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
NYRR Dash	p < 0.0001	p < 0.0001	p = 0.2007	p = 0.5261	p < 0.0001
OC	p < 0.0001	p < 0.0001	p = 0.0102	p = 0.0161	p = 0.0035
Percy Sutton	p = 0.0210	p = 0.0285	p < 0.0001	p < 0.0001	p < 0.0001

Source: (Boston 5K, 2013; Chelsea's Run 5K, 2013; Gasparilla 5K, 2013; McGuire's 5K, 2013; MLB All Star 5K, 2013; NYRR Dash 5K, 2013; OC 5K, 2013; Percy Sutton 5K, 2013)

Descriptive statistics for ages based on the 8 five kilometer races in the sample are given in Table 32. Test results are given in Table 33 for the ages.

The mean ages for the male and female finishers were not significantly different for the OC five kilometer race and Chelsea's run. The mean age for the male finishers was significantly larger than the mean age for the female finishers in 5 of the races and marginally significantly larger in 1 of the races. The median age for the male and female finishers were not significantly different for the McGuire's St Patrick's Day, OC, and Chelsea's Run five kilometer race. The median age for males was significantly larger than the median age for females in 5 of the 5K races. The variance of the ages for males was significantly higher than females for all of the 5K races. The distributions of ages between males and females was significantly different in 6 of the 5K races.

Table-32. Descriptive Statistics for Age – Five Kilometer

Five K Race	Mean		Median		Variance		Standard Dev	
Boston	M: 39.4	F: 36.9	M: 38.0	F: 35.0	M: 154.5	F: 129.1	M: 12.43	F: 11.36
Chelsea's Run	M: 36.2	F: 35.1	M: 37.0	F: 35.0	M: 216.6	F: 171.8	M: 14.72	F: 13.11
Gasparilla	M: 40.3	F: 38.5	M: 40.0	F: 37.0	M: 201.2	F: 166.6	M: 14.19	F: 12.91
McGuire's	M: 34.3	F: 33.9	M: 31.0	F: 31.0	M: 154.4	F: 138.9	M: 12.43	F: 11.79
MLB All Star	M: 35.5	F: 33.2	M: 34.0	F: 32.0	M: 101.1	F: 77.8	M: 10.06	F: 8.82
NYRR Dash	M: 39.5	F: 36.9	M: 38.0	F: 35.0	M: 133.9	F: 115.0	M: 11.57	F: 10.73
OC	M: 37.0	F: 37.0	M: 36.0	F: 36.0	M: 173.1	F: 170.1	M: 13.16	F: 13.04
Percy Sutton	M: 37.6	F: 34.6	M: 36.0	F: 32.0	M: 124.6	F: 104.9	M: 11.17	F: 10.24

Source: (Boston 5K, 2013; Chelsea's Run 5K, 2013; Gasparilla 5K, 2013; McGuire's 5K, 2013; MLB All Star 5K, 2013; NYRR Dash 5K, 2013; OC 5K, 2013; Percy Sutton 5K, 2013)

Table-33. Test Results for Age – Five Kilometer

Five K Race	F Test	Levene's Test	T Test	Wilcoxon	KS Test
Boston	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
Chelsea's Run	p = 0.0011	p < 0.0001	p = 0.1280	p = 0.1753	p = 0.0461
Gasparilla	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
McGuire's	p < 0.0001	p < 0.0001	p = 0.0911	p = 0.3831	p = 0.2264
MLB All Star	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
NYRR Dash	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001	p < 0.0001
OC	p = 0.7950	p = 0.7890	p = 0.9653	p = 0.7642	p = 0.9905
Percy Sutton	p = 0.0003	p = 0.0044	p < 0.0001	p < 0.0001	p < 0.0001

Source: (Boston 5K, 2013; Chelsea's Run 5K, 2013; Gasparilla 5K, 2013; McGuire's 5K, 2013; MLB All Star 5K, 2013; NYRR Dash 5K, 2013; OC 5K, 2013; Percy Sutton 5K, 2013)

6. META-ANALYSIS RESULTS

Having performed statistical tests to find differences between male and female race finishers for each race individually, we wanted to combine the results in order to draw conclusions for each race distance. To accomplish this task, we converted the two-sided p-values into one-sided p-values that answer the alternative hypothesis of male results larger than female results; the test for proportions of less than or equal to forty percent was also converted to answer the alternative hypothesis of female results larger than male results; these conversions were performed for every statistical test. After the p-values were converted, we performed a meta-analysis to combine the p-values of the three qualifying marathons using Stouffer's weighted Z-score method for combining the p-values as described in section 2.1. For ease of computation, we used a p-value of 0.00009 for all p-values less than 0.0001 and a p-value of 0.99999 for all p-values equal to 1.000 that appeared in the test results.

Table-34. Meta-Analysis for Marathon Results

Variable	Age/Gender Graded Scores		Ages	
	Z value	P value: M > F	Z value	P value: M > F
Proportions				
60%	5.678	p < 0.0001	NA	NA
70%	6.687	p < 0.0001	NA	NA
40%*	-1.106	p = 0.8656	NA	NA
Variances				
F Test	7.545	p < 0.0001	8.533	p < 0.0001
Levene's Test	7.558	p < 0.0001	8.585	p < 0.0001
Means	5.158	p < 0.0001	8.626	p < 0.0001
Medians	4.501	p < 0.0001	8.626	p < 0.0001

*40% proportions were tested for males less than females

The results for all non-qualifying marathons included in this paper were combined in order to determine if the individual hypotheses held for the larger population of marathon finishers; we used the combined data for the eight non-major marathons to examine the variables age/gender graded score and age of race finisher. For the variable age/gender graded score, there was no statistical evidence that the proportion of race finishers with an age graded score of forty percent or lower for female runners was larger than the proportion of race finishers with an age/gender graded score of forty percent or lower for male runners. All other tests produce highly significant results with p-values less than 0.0001 which means that the mean, median, and variance of age/gender graded scores and ages for male finishers were significantly larger than the mean, median, and variance of age/gender graded scores and ages for female finishers. We can also say that the distributions of age/gender graded scores and ages for male finishers was significantly different than the age graded scores for female finishers. The results are given in Table 34.

The results of the meta-analysis for the combined half marathon finishers are all highly significant with p-values less than 0.0001. From the results we can easily see that the proportions of finishers with an age/gender graded score of sixty percent or higher and seventy percent or higher for male runners were significantly higher than the proportion of finishers with an age/gender graded score of sixty percent or higher and seventy percent or higher for female runners.

Table-35. Meta-Analysis for Ten Kilometer Results

Variable	Age/Gender Graded Scores		Ages	
	Z value	P value: M > F	Z value	P value: M > F
Proportions				
60%	7.757	p < 0.0001	NA	NA
70%	7.586	p < 0.0001	NA	NA
40%*	6.594	p < 0.0001	NA	NA
Variances				
F Test	7.343	p < 0.0001	7.598	p < 0.0001
Levene's Test	7.329	p < 0.0001	7.585	p < 0.0001
Means	7.757	p < 0.0001	7.289	p < 0.0001
Medians	7.757	p < 0.0001	7.601	p < 0.0001

*40% proportions were tested for males less than females

For the proportion of finishers with an age/gender graded score of forty percent or lower, the proportion of female finishers was significantly higher than the proportion of male finishers. The mean, median, and variance of age/gender graded scores and ages for male finishers were significantly larger than the mean, median, and variance of age/gender graded scores and ages for female finishers.

The results for all ten kilometer races included in this paper were combined in order to determine if the individual hypotheses held for the larger population of ten kilometer finishers. The results of the meta-analysis for the combined ten kilometer race finishers were all highly significant with p-values less than 0.0001. From the results we can easily see that the proportion of finishers with an age/gender graded score of sixty percent or higher and seventy percent or higher for male runners were significantly higher than the proportion of finishers with an age/gender graded score of sixty percent or higher and seventy percent or higher for female runners. For the proportion of finishers with an age/gender graded score of forty percent or lower, the proportion of female finishers is significantly higher than the proportion of male finishers. We can also see that the mean, median, and variance of age/gender graded scores and ages for male finishers were significantly larger than the mean, median, and variance of age/gender graded scores and ages for female finishers. Results are given in Table 35.

The results for all five kilometer races included in this paper were combined in order to determine if the individual hypotheses held for the larger population of five kilometer finishers. The results of the meta-analysis for the combined five kilometer race finishers were all highly significant with p-values less than 0.0001. The proportion of finishers with an age graded score of sixty percent or higher and seventy percent or higher for male runners was significantly higher than the proportion of finishers with an age graded score of sixty percent or higher and seventy percent or higher for female runners. For the proportion of finishers with an age graded score of forty percent or lower, the proportion of female finishers was significantly higher than the proportion of male finishers. The mean, median, and variance of age graded scores and ages for male finishers were significantly larger than the mean, median, and variance of age graded scores and ages for female finishers.

7. CONCLUSIONS

The goal of this paper was to examine the differences between the male and female running populations in order to determine if these differences corroborate the theory that men are more competitive than women. If this theory is true we would expect to see certain trends throughout our data; specifically, we would see the age/gender graded scores for men to be higher and closer together than the age/gender graded scores for women. Translated into numbers, we would expect to see a larger proportion of male finishers with an age/gender graded score at or above sixty percent as well as a larger proportion of female finishers with an age/gender graded score at or below forty percent. The mean and median age/gender graded score for men would be larger than the scores for women while the variance of the age/gender graded scores for women would be larger than the variance of the male finishers, if the male finishers are indeed more competitive.

To determine if these trends appear in the populations of male and female runners, we conducted several statistical tests for the variables of age/gender graded score and age of the runners; we preformed this analysis for three marathons that require qualification to participate and thirty-two races that did not require qualification to participate, eight races at each of the four popular race distances: marathon, half marathon, ten kilometer, and five kilometer. Specifically, we tested the significance of the differences between the male and female running populations for the proportion of race finishers with an age/gender graded score of sixty percent or higher, the proportion of race finishers with an age/gender graded score of forty percent or lower, the mean and median age/gender graded scores, the variance of the age/gender graded scores, and the distribution of the age/gender graded scores. We also tested the mean and median ages, the variance of the ages, and the distribution of ages.

Overall, we found that across all distances, the mean age/gender graded score for male race finishers was significantly larger than the mean age/gender graded score for female race finishers; this shows that on average, the male runners finished their chosen race distance faster than the female runners in the same race. This finding is echoed in the median age/gender graded score which is significantly larger for the male runners than for the female runners; showing that, on average, the male racers finished their chosen race distance faster than the female runners

in the same race. These findings agree with what we would expect to see if men were more competitive than women; our next result does not agree with what we would expect.

We found that, for all distances examined, the variance of the age/gender graded scores for the male runners is significantly larger than the variance of the age/gender graded scores for the female runners; this is for the races and distances overall, with some specific races may have a larger variance for the female population than for the male population. This finding suggests that the male runners are possibly not more competitive than the female runners. It suggests that that the age/gender graded scores of male runners are more spread out, with some male runners doing very well and some not as well. The age/gender graded scores of female runners were bunched closer together.

To further examine which gender was more competitive, we looked at the proportion of finishers at certain levels of the age/gender graded scores to determine which gender had a larger proportion of race finishers at that level. For all distances that we examined, we found that the proportion of male finishers with an age/gender graded score of sixty percent or higher is significantly larger than the proportion of female finishers with an age graded score of sixty percent or higher. This result shows that, for all distances, the male running population had a higher proportion finishing comparatively faster than the female running population which suggests that the male runners are more competitive than the female runners. The results for the proportion of race finishers at or below forty percent are less straight forward. For competitive marathons studied, we found that the proportion of male runners with an age/gender graded score of forty percent or lower is significantly larger than the proportion of female runners with an age/gender graded score of forty percent or lower. In other marathons, there was no significant difference, but the sample proportion of women was lower. For all other race distances studied, the proportion of female finishers with an age/gender graded score of forty percent or lower is significantly larger than the proportion of male finishers with an age/gender graded score of forty percent or lower. These findings show that more men finish comparatively faster than women and more women finish comparatively slower than men for half marathon, ten kilometer, and five kilometer races which suggest that men are more competitive than women for these race distances. The marathon race distance results show that more men finish comparatively faster than women but more men also finish comparatively slower than women.

REFERENCES

Airforce Marathon, 2013. Available from http://www.usafmarathon.com/ [Accessed May, 2015].

Berkley Half-Marathon, 2013. Available from http://berkeleyhalf.com/ [Accessed May, 2015].

Booth, A. and P. Nolen, 2011. Choosing to compete: How different are girls and boys? Journal of Economic Behavior & Organization, 81(2): 542-555.

Boston 5K, 2013. Available from http://www.baa.org/races/5k.aspx [Accessed May, 2015].

Boston Marathon, 2012. Available from http://www.baa.org/races/boston-marathon.aspx [Accessed May, 2015].

Brooklyn Half-Marathon, 2013. Available from http://web2.nyrrc.org/cgi-bin/start.cgi/aes-programs/results/startup.html?result.id=b30518&result.year=2013 [Accessed May, 2015].

Burfoot Amby, 2007. The history of the Marathon. Sports Medicine, 37(4-5): 284-287.

Capital City Half-Marathon, 2013. Available from http://www.capitalcityhalfmarathon.com/ [Accessed May, 2015].

Cashdan, E., 1998. Are men more competitive than women? British Journal of Social Psychology, 37(2): 213-229.

Chelsea's Run 5K, 2013. Available from https://chelseaslight.org/programs/finish-chelseas-run/ [Accessed May,2015].

Chicago Marathon, 2013. Available from http://www.chicagomarathon.com/ [Accessed May, 2015].

Columbus Half-Marathon, 2013. Available from http://www.columbusmarathon.com/ [Accessed May, 2015].

Cooper River 10K, 2013. Available from http://web2.nyrrc.org/cgi-bin/start.cgi/aes-programs/results/startup.html?result.id=b30317&result.year=2013 [Accessed May, 2015].

Crescent City 10K, 2013. Available from http://www.ccc10k.com/ [Accessed May, 2015].

Dato, S. and P. Nieken, 2014. Gender differences in competition and sabotage. Journal of Economic Behavior & Organization, 100(1): 64-80.

Duluth Half-Marathon, 2013. Available from http://www.grandmasmarathon.com/racesevents/garry-bjorklund-half-marathon/register-garry-bjorklund-half-marathon/ [Accessed May, 2015].

Fargo 10K, 2013. Available from http://www.fargomarathon.com/ [Accessed May, 2015].

Fisher, R.A., 1924. On a distribution yielding the error functions of several well known statistics. Proceedings of the International Congress of Mathematics, Toronto, 2: 805-813.

Frick, B., 2011. Gender differences in competitiveness: Empirical evidence from professional distance running. Labour Economics, 18(3): 389-398.

Gasparilla 5K, 2013. Available from http://www.tampabayrun.com/5k-general-information [Accessed May, 2015].

Grandma's Marathon, 2013. Available from http://www.grandmasmarathon.com/racesevents/grandmas-marathon/register/ [Accessed May, 2015].

International Association of Athletics Federations, 2015. Olympics games - history. IAAF: Archive of Past Events. Available from www.iaaf.org.

International Olympic Committee, 2015. Past games – summer olympics. Olympic Games: Past Games. Available from www.olympics.org.

Joe Kleinerman 10K, 2013. Available from http://web2.nyrrc.org/cgi-bin/start.cgi/aes-programs/results/startup.html?result.id=b30105&result.year=2013 [Accessed May, 2015].

Kolmogorov, A.N., 1933. Sulla determinazione empirica di una legge di distribuzione. Giornalle Dell'Instituto Italiano Degli Attuari, 4(1): 83- 91.

Levene, H., 1960. Robust tests for equality of variances. Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling, 2(Chapter 25): 278-292.

Los Angeles Marathon, 2013. Available from http://www.lamarathon.com/ [Accessed May, 2015].

Lovett, C.C., 1997. Olympic Marathon: A centennial history of the games most storied race. Westport, CT: Praeger Publishers.

Manhattan Beach 10K, 2013. Available from http://mb10k.com/ [Accessed May, 2015].

Manhattan Half-Marathon, 2013. Available from http://web2.nyrrc.org/cgi-bin/start.cgi/aes-programs/results/startup.html?result.id=b30127&result.year=2013 [Accessed May, 2015].

Mann, H.B. and D.R. Whitney, 1947. On a test of whether one of two random variables is stochastically larger than the other. Annals of Mathematical Statistics, 18(1): 50-60.

McGuire's 5K, 2013. Available from http://www.mcguiresirishpub.com/race-results.php [Accessed May, 2015].

MLB All Star 5K, 2013. Available from http://web2.nyrrc.org/cgi-bin/start.cgi/aes-programs/results/startup.html?result.id=b30713&result.year=2013 [Accessed May, 2015].

New York City Half-Marathon, 2013. Available from http://web2.nyrrc.org/cgi-bin/start.cgi/aes-programs/results/startup.html?result.id=b30317&result.year=2013 [Accessed May, 2015].

New York Marathon, 2013. Available from http://www.nyrr.org/races-and-events/2013/the-2013-ing-new-york-city-marathon [Accessed May, 2015].

NYRR Dash 5K, 2013. Available from http://web2.nyrrc.org/cgi-bin/start.cgi/aes-programs/results/startup.html?result.id=b30713&result.year=2013 [Accessed May, 2015].

OC 5K, 2013. Wahoo's OC 5K. Available from http://www.ocmarathon.com/races/oc-5k/wahoos-oc-5k.aspx.

Ogden Marathon, 2015. Available from http://getoutandlive.org/participate/ogden-marathon [Accessed May, 2015].

Pate, R.R. and J.R. O'Neill, 2007. American women in the Marathon. Sports Medicine, 37(4-5): 294-298.

Percy Sutton 5K, 2013. Available from http://web2.nyrrc.org/cgi-bin/htmlos.cgi/2147.1.215572612383251676 [Accessed May, 2015].

Philadelphia Marathon, 2013. Available from http://philadelphiamarathon.com/ [Accessed May, 2015].

Queens 10K, 2013. Available from http://web2.nyrrc.org/cgi-bin/start.cgi/aes-programs/results/startup.html?result.id=b30721&result.year=2013 [Accessed May, 2015].

Road Runners Club of America, 2015. FUN damentals of youth running. Road Runners Club of America: Education and Advocacy. Available from www.rrca.org.

Runner's World, 2013. Boston Marathon and New York Marathon. Runner's World: Races and Places. Available from www.runnersworld.com.

Running USA, 2014. State of the sport – part III US road trends. Running USA: Annual Reports.

San Francisco Marathon, 2013. Available from https://www.thesfmarathon.com/ [Accessed May, 2015].

Schultz, J., 2015. Going the distance: The road to the 1984 olympic women's Marathon. International Journal of the History of Sport, 32(1): 72-88.

Scotland Run 10K, 2013. Available from http://web2.nyrrc.org/cgi-bin/start.cgi/aes-programs/results/startup.html?result.id=b30406&result.year=2013 [Accessed May, 2015].

Smirnov, N., 1948. Table for estimating the goodness of fit of empirical distributions. Annals of Mathematical Statistics, 19(2): 279-281.

Stouffer, S.A., S.A. Edward, D.C. Leland, S.A. Shirley and W.J.M. Robin, 1949. The American soldier: Adjustment during army life. (Studies in Social Psychology in World War II). Princeton University Press, 1.

Surf City Marathon and Half-Marathon, 2013. Available from http://www.runsurfcity.com/ [Accessed May, 2015].

Thibault, V., 2010. Women and men in sport performance: The gender gap has not evolved since 1983. Journal of Sports Science & Medicine, 9(2): 214 – 223.

Twin Cities Marathon, 2013. Available from https://www.tcmevents.org/events/medtronic_twin_cities_marathon_weekend_-_october_2-4_2015/marathon/ [Accessed May, 2015].

UAE Healthy Kidney 10K, 2013. Available from http://web2.nyrrc.org/cgi-bin/start.cgi/aes-programs/results/startup.html?result.id=b30511&result.year=2013 [Accessed May, 2015].

Warner, S. and M.A. Dixon, 2015. Competition, gender and the sport experience: An exploration among college athletes. Sport, Education and Society, 20(4): 527-545.

Weiss, N., 2012. Introductory statistics. 9th Edn., Boston, MA: Pearson Publishing.

Wilcoxon, F., 1945. Individual comparisons by ranking methods. Biometrics Bulletin, 1(6): 80-83.

WMA, 2010. Road age standards – WMA 2010. World Master's Athletics: Laws & Rules. Available from www.world-masters-athletics.org.

World Master's Athletics, 2010. Road age standards – WMA 2010. World Master's Athletics: Laws & Rules. Available from www.world-masters-athletics.org.