

Changes to Bill Page Methodology for 2012

1. Introduction

Ever since the Bill Page Memorial Trophies were first instituted in 1993 the methodology has remained unchanged. The age/sex factors used are those produced in 1994 by WAVA; the resulting age/sex-graded scores for the three relevant races have been added together to produce a total, the highest winning.

2. Issues with Current Methodology

- The 1994 factors are out of date
- The 1994 factors are track-based up to and including 10k
- The adding together of the three scores is arithmetically invalid

The first issue is resolved by moving to the latest factors produced by WMA, which incorporate a further sixteen years of analysis of the relative performances of veteran and senior athletes. They are called the 2010 factors, see: <http://home.roadrunner.com/~alanjones/AgeGrade.html> The validity of the 2010 factors is reinforced by their planned inclusion in the next release of *RaceMaster*, which will also have the option of calculating age/graded scores on either Gun or Chip time.

The second issue is resolved through WMA having produced the new factors in two versions: track; road running. The road running factors start at 5k and thus allow road-based factors to be used for the first time in scoring all SGP events.

The third issue came to light when Haywards Heath Harriers developed the methodology for Cooper Cup Award for the Southdowns Way Relay. This is awarded to the team with the best age/sex-graded score, which required a methodology that took into account the age and sex of each team member and the fact there are eighteen legs in the race.

3. Two Ways to Score

The difference between what has been done in the Bill Page Award so far and what is considered is more valid method is best illustrated by an example.

Consider an event where the world record is 240 seconds and a 45 year old athlete runs 360 seconds. Assume the age factor for this event for a 45 year old is 0.80

At the moment we calculate the world standard of a 45 year old for the event to be $240/0.80 = 300$ seconds. We then calculate this athlete's score as $300/360 = 83.33\%$

Alternatively, we calculate the senior-equivalent time to be $360*0.80 = 288$ and express the world record as percentage of it, giving the identical score: $240/288 = 83.33\%$

More generally, if f is the age factor, T the time and W the world record, then the score is calculated as: $(W/f)/T$ or $W/(fT)$, both giving 83.33% for the example above

But why express the world record (240 seconds) as a percentage of the 45 year-old's senior-equivalent time (288 seconds)? Perhaps this is the done as it is felt more intuitive to aim for 100% from below. It is more meaningful to express the athlete's senior-equivalent time as a percentage of the world record, giving $288/240 = 120\%$.

In other words, $(fT)/W$, and thus W , the denominator is fixed for all competitors.

Under the old method the highest score wins – under the new method the lowest score wins. Clearly, for one race whichever method is used the respective age/sex-graded rankings will be identical.

So, for one race either method can be used – and perhaps the old method does have a *presentational* advantage in that it produces scores that are below 100. However, difficulties arise when more than one race is considered – as with the Bill Page Award – and this is when it can be argued that the second method of scoring is additive by virtue of all runners in a given race having their senior-equivalent times expressed as a percentage of the same denominator, namely the world record. This is what allows scores produced by the new method to be added.

4. Methodology for 2012

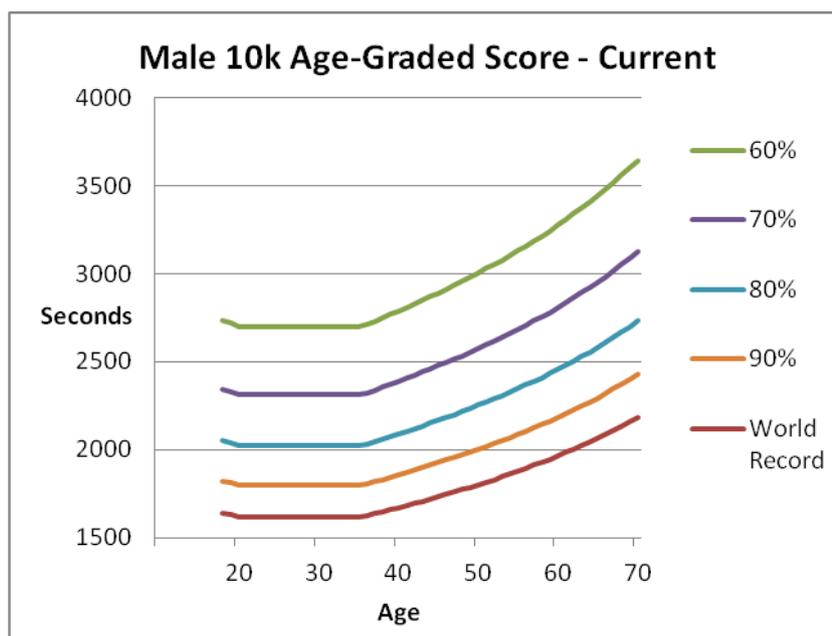
The WMA’s 2010 road-based factors will be used and scores for individual races will still be calculated and presented under the present methodology. However, when producing the total of an individual’s best three scores, the calculation will be made by adding the scores derived under the new method, and then for *presentational* purposes, the total will be converted and scaled to give a figure comparable with those that have been produced in the past.

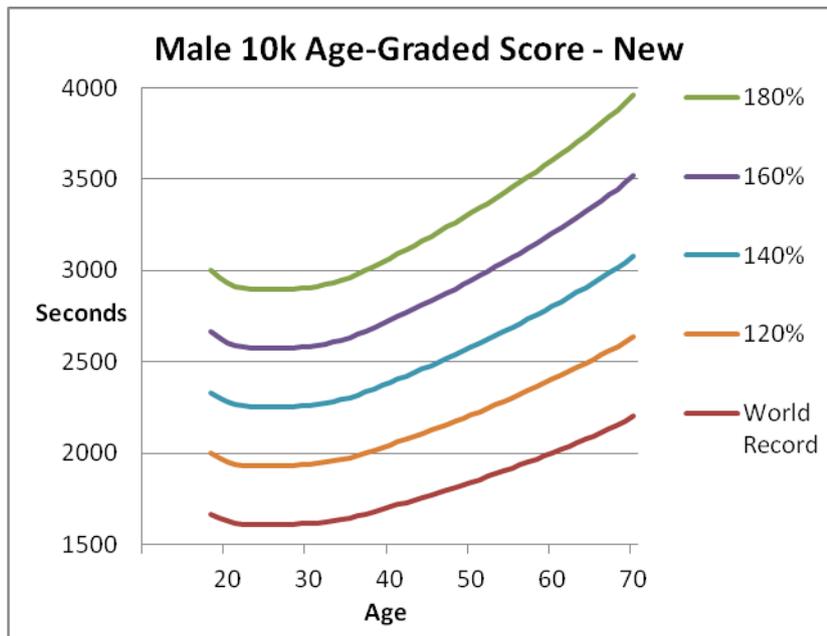
Example: a runner’s three best scores are: 72; 76; 80.

Under the current method the total Bill Page result is obtained by calculating “3 times the average score”. The average of the three scores is 76, giving $3 \times 76 = \mathbf{228.00}$

Under the new method the scores are: 138.89, 131.58, 125.00 giving an average of 131.82. The average is now converted to the old basis as $1/131.82 \times 10000 = 75.86$. It is then multiplied by 3, giving **227.58**

Although the difference in this particular example is small, the ranking of a group of runners under both the current and new methods may well be different; in other words, the first-placed runner will not *necessarily* be the same in both cases. So which method of combining the three scores gives the truest measure of “best”? Consider the following complementary perspective. Under the current method an athlete’s improvement in time does not give a linear improvement in percentage score. A 40 year old male whose 10k time is 47:21 (2841 seconds, giving a score of 60%) finds that to improve to 70% he needs to reduce his race time by 406 seconds; however, to go from 70% to 80% he only needs to shave off 304 seconds. Under the new method there is a linear relationship. The time of 2841 seconds scores 167% and every 10% reduction is achieved by a constant 170 second race-time reduction. This is illustrated in the charts below.





For the new method the lines are equally spaced, illustrating the linearity discussed above.

5. Comparability Issues

Comparison of the aggregate winner's score for the Bill Page trophy over time has to be done with great care. For instance:

- Results for 2008 and earlier years were based on scores from three nominated races of distance 10k, 10 miles and half marathon. The nominations changed over the years making it difficult to compare performances over time. Although the 10k has always been Brighton or Newick, the 10 mile has been variously: Worthing, Great South Run, Hailsham and the off-road Lewes Downland. The half marathon has moved from Brighton to Worthing to Barns Green back to Brighton and back to Barns Green.
- In addition, in 1997 and 2002 a 10k race was substituted for the 10 mile event
- For 2009 the criterion was changed from three specific races to: "three best scores in SGP events, at least one of which must be 10 miles or longer". Thus a discontinuity was introduced. In addition a decision was made to use Chip Time, giving rise to a further discontinuity.
- Before chip timing was available extreme "bottlenecking" at larger races (eg Brighton 10k, Barns Green half marathon) had a major impact on age-graded performances
- For at least one past year the effect of moving from 1994 factors to 2010 factors – and leaving unchanged the methodology for adding scores – changes the winner of the Bill Page Award
- Results for 2012 will not be comparable with those from 2009 to 2011 because of the new factors and the new method for adding scores

It is clear that the original 1994 factors were much too stringent in their expectations of the performance levels of older female runners. For instance, a 70 year old female running 53:13 for 10k would have scored 80.00% under the 1994 factors; however, with the new factors she would score 85.74%. How many older female Harriers were denied their due award, and indeed satisfaction, of outscoring the male winner? We shall never know. It is clear that this must be corrected.

The chart below shows how the times necessary to achieve a score of 80% over 10k have increased through moving from the 1994 factors to those for 2010. For men the maximum difference is 51 seconds at age 55. For women the difference is ever increasing from the age of 46, reaching 229 seconds at age 70.

