ABSTRACT
Every year over 40 000 runners complete Gothenburg Half Marathon, one of the world’s largest half-marathons. As participation in recreational races become more common among e.g., older people and those without extensive training experience, providing advice to recreational runners on how to plan the pacing during race is valuable for a safer, more positive experience, lessening the risk of over-straining, injury, or collapse. We conduct a large-scale data analysis of 10 years (2011 – 2019) of publicly available results data (n = 423 496), investigating which runners pace well and which ones experience slowdowns (“hitting the wall”), depending on age, sex, ability, and temperature on the race day. Most runners loose several minutes on the second half and may have benefited from pacing advice. Men are about twice as likely to hit the wall compared to women, and younger runners at more likely to do so than middle-aged. Increased temperature led to both slower finishing times and, among men, also an increased risk of hitting the wall. Our results can be used by race organisers to provide advice to participants based on e.g., the weather prognosis on the race day, as well as estimating need for medical assistance.

KEYWORDS
Half-marathon, running, pacing patterns, recreational running, results data

1. Introduction

Gothenburg Half Marathon is one of the worlds largest half-marathons. Most participants are recreational runners of all ages and fitness levels, and many return to participate many times. An active lifestyle, including recreational running can contribute to public health (Lee et al., 2017), why it is of interest to support runners to pace well for a pleasant experience, encouraging return participation while avoiding injuries, hitting the wall or in extreme cases collapsing. Our goal is to investigate if we can use a large database of historic public results to analyse what is indicative of both good and bad pacing performance: which runners manage to complete the race without losing time on the second half, and which runners experience dramatic slowdown? We base our work on public results data for ten years (2010 – 2019) from Gothenburg Half Marathon (n = 423 496) where finishing times and 5 km split times are recorded. We follow a similar methodology as Smyth (2021) where “hitting the wall” is defined as a > 25% drop in 5 km split time.
We find that there is room for improvement in pacing among recreational runners: most runners lose quite a lot of time on the second half of the race. Men are twice as likely as women to hit the wall, and runners below the age of 30 are more likely to hit the wall than middle-aged runners. However, among the younger runners there is also a larger share who manage to pace well and not lose time on the latter part of the race. Ability also plays a role, the proportion of runners hitting the wall increase with increased finishing time, as one would expect, while the converse is true for those running at a more even pace. Finish time and pacing is negatively affected by increasing temperatures, especially among men there is an increased risk of hitting the wall.

Our work covers over 400 000 race records, all from the same race over a 10-year period, which to our knowledge is one of the largest investigations of pacing patterns for a half-marathon, and also allow us to compare results for the same race from years with temperatures ranging from 13-25°C. The result can be used by race organisers to inform participants of risk factors affecting desired running pace. In the longer run these results could potentially also feed into the development of, for example, a personalised pacing app helping recreational participants finding a pace suitable for their own fitness level and the conditions on the day of the race.

**Related Work**

Several works investigate pacing patterns on the marathon distance, including the risk of large slowdowns based on e.g., age, sex, and ability (Berndsen, Lawlor, & Smyth, 2020; Deaner, Carter, Joyner, & Hunter, 2015; March, Vanderburgh, Titlebaum, & Hoops, 2011; Smyth, 2021). Results points toward younger and male runners being more at risk of slowdowns, while women and older runners generally pacing more evenly. Similar patterns were also found in an analysis of the Vienna half-marathon in 2017 (Čuk, Nikolaidis, & Knechtle, 2020; Čuk, Nikolaidis, Markovic, & Knechtle, 2019). Nikolaidis, Čuk, and Knechtle (2019) found that half-marathon runners pace more evenly than marathon runners, but still generally adopt a positive pacing strategy, gradually slowing down during the race. Knechtle and Nikolaidis (2018) investigated age differences in finishing times on Gothenburg Half Marathon between 2014-2016 and found the fastest finishing times in women aged below 40, and men between 35-39.

Ely, Cheuvront, Roberts, and Montain (2007) investigated the impact of weather and temperature on marathon finishing times, and found trends towards slowing with increased wet-bulb globe temperature. Trubee, Vanderburgh, Diestelkamp, and Jackson (2014) finds that for non-elite marathon runners, women pace better than men, and that this is magnified in hotter temperatures. Gothenburg Half Marathon has been the subject of several previous studies investigating the incidence and characteristics of runners collapsing or requiring medical assistance, showing a higher incidence in warm years, and among younger runners (Carlström et al., 2019; Khorrarn-Manesh et al., 2020; Lüning, Mangelus, Carlström, Nilson, & Börjesson, 2019). Similar studies have also been undertaken for half-marathons in South Africa (Schwabe, Schwellnus, Derman, Swanevelder, & Jordaan, 2014a, 2014b). As opposed to these studies, we here instead look at pacing patterns for much the large group of runners who complete the race, and how they paced themselves between splits.
2. Materials and Methods

2.1. Data

Our data consists of results from Gothenburg Half Marathon from the years 2010 – 2019 (earlier years did not have split times available). This data is publicly available from the race organisers website\(^1\), we work with a snapshot of the underlying results database retrieved on 2 November 2021. Of relevance to analysis is the finish time, split times at 5, 10, 15 and 20 km, year of birth and sex. Each unique runner is identified by a unique numeric ID. In addition, we also added information about the temperature on the race day each year, obtained from the Swedish Meteorological and Hydrological Institute. As the runners start in different groups throughout the afternoon, we simply used the temperature at 3pm as actual temperature is assumed to be similar. The average daytime top temperature for Gothenburg in the month of May (when the race is held) is 17°C.

After pre-processing to removing entries with missing or obviously faulty information (e.g. missing/incorrect split- and finishing times) we obtain a dataset of 423 496 records (female = 140 409; male = 283 087). As there are repeat participants, the dataset contains 184 890 unique individuals, on average participating 2.3 times in the ten-year period. Table 1 summarises the data by year.

2.2. Methods

Our aim is to estimate which recreational runners manage to pace well, and which pace badly. While optimal pacing depend on many factors (Roelands, de Koning, Foster, Hettinga, & Meeusen, 2013), we will here not consider optimising placement or finishing time, but rather completing the race comfortably, with low risk of collapsing or injury. From a broader health perspective, this could make the experience more positive, encouraging repeat participation. Recreational half-marathon runners are commonly advised to pace themselves evenly throughout the race, but often in fact slow down throughout (Nikolaidis et al., 2019). Whether or not an even pace is really the optimal pacing strategy for half marathon is somewhat unclear (Abbiss & Laursen, 2008), however, for our purposes, counting the number of runners managing a negative or even half-way split is a good enough proxy. Runners pacing in this manner will are unlikely to have overextend themselves, and are at low risk of medical complications. On the other hand, we estimate bad pacing, "hitting the wall", to be those who somewhere experience a dramatic slowdown, with a severe drop in pace during some 5 km segment. Even though hitting the wall is not a precise measure of the physiological state, there is not even a generally agreed upon precise definition (Berndsen et al., 2020; Buman, Oml, Giacobbi, & Brewer, 2008), it serves as an estimate of overextension which is undesirable: such runners may be at risk of collapsing, and will probably have a less enjoyable experience.

To identify and compare pacing patterns we use two metrics defined below: the \textit{Split Difference (SD)} capturing time gained or lost during the second half of the race, and the \textit{Degree of Slowdown (DoS)} to identify any drastic slow downs during any 5 km split.

\(^1\)https://reg.goteborgsvarvet.se/sok/resultatlista.aspx
**Split Difference.** The Split Difference captures time lost/gained compared to maintaining the same pace held at the 10 km mark for the rest of the race. As no exact mid-point split is available, we define $SD$ as:

$$SD = \text{FinishTime} - 10\text{kmSplit} \times 2.109775 \quad (1)$$

$SD < 0$ indicate that the runner was faster on the second half (a negative split), while $SD > 0$ indicate they slowed down (a positive split). Only 9.8% of runners overall manage to run an even or negative split.

**Hitting the Wall.** We apply the operational definition by Smyth (2021) where a 25% drop in pace between two splits is used as a proxy for having hit the wall. The base-pace ($BP$) is defined as the average pace over the 5 and 10 km splits, as runners may experience congestion at the start of the race prohibiting them from directly matching the planned pace. Further, the risk of hitting the wall this early in the race is low. The $BP$ is then compared to the degree of slowdown (DoS) for a segment $s$ (10-15 km, 15-20 km or 20-21 km), defined as the ratio of segment pace and base pace:

$$DoS(s) = \frac{\text{pace}(s) - BP}{BP} = \frac{\text{pace}(s)}{BP} - 1 \quad (2)$$

Overall, 8.6% of runners in our dataset have hit the wall on some segment using this definition, most commonly between 15-20 km. As expected, this is a smaller proportion than in studies on marathon, as slowdowns during half-marathons more likely is due to lactate buildup or simply fatigue from overextension during the first half, rather than glycogen depletion.

### 2.3. Data Analysis

We use a Python scipy.stats library for statistical analysis. For pairwise comparisons we use a Fisher Exact test, and for multiple groups a chi-square test. Python scripts are available on request.

### 3. Results

#### 3.1. Pacing Patterns

The average runner in our analysis start the race at a faster pace than they can maintain, and gradually slow down by each 5 km split, until the 20 km mark, when they manage an “end spurt” and increase their speed when the goal is in sight (Figure 1a). Runners who hit the wall display the same pacing pattern, but with an even faster start and larger drop in pace between 10-20 km. Runners who on the other hand manage a negative split, take it easier in the first 5 km, to then maintain a very even pace, until the final end-spurt. Consequently, most runners also loose time in the second half of the race (Figure 1b). The fastest runners, finishing in less than 90 minutes, loose on average around 1:30 minutes on the second half, while an average runner, finishing in 120 minutes loose just over 4 minutes. Note that among the very fastest, it seems few run a negative split, possibly due to race tactics. In the groups with
slower finishing times there is much more spread possible among the pacing strategies, but on average, slower finishers also lose more on the second half.

Average split differences are very similar between men and women, however grouping runners by finishing time shows the women consistently lose slightly less time for runners with finishing time > 90 minutes (see Supplementary Material Figure 5).

Older runners (50+) have a higher average finish time and split difference. However, grouping by finishing time shows no differences except among the slower runners (finish time > 150 minutes), where the younger age-groups in fact lose more time (see Supplementary Material Figure 6).

3.2. Who Pace Well and Who Hits the Wall?

Next, we investigate which runners, bases on sex, age, and ability, manage a good pacing strategy (achieving a negative or equal split), and conversely, which runners end up hitting the wall.

3.2.1. Sex

Men are twice as likely to hit the wall: 10.2% of male runners do so compared to just 5.4% of women (OR = 2.0; p < 0.001), see Table 1. Most runners will slow down during the second half of the race (see Figure 1), but among runners managing a negative or equal split, male and female runners perform similarly: 10.1% of male runners and 9.2% of female do so (OR = 1.1; p < 0.001).

3.2.2. Age

Gothenburg Half marathon is open for participants aged 17 and above, with most runners being between 30-49 years old, see Table 2. Age information was missing or incorrect for 3173 datapoints, which were excluded from analysis.

The increased risk of hitting the wall for men compared to women is consistently high across all age groups (1.74 ≤ OR ≤ 2.29; p < 0.001), see Figure 2a. Both among women and men the youngest runners are most likely to hit the wall, while the 40–49-year-olds are least likely (female: OR = 0.59; male: OR = 0.66; p < 0.001).
between consecutive age groups within sex are significant except for the women in their 50’s vs. 60’s ($p = 0.74$).

A larger proportion of younger participants run a negative split, with men slightly higher than females consistently across age groups ($1.17 \leq OR \leq 1.49; p < 0.001$). This decreases for each older age group, see Figure 2b (pairwise between consecutive age groups, women: $0.58 \leq OR \leq 0.80$; men: $0.68 \leq OR \leq 0.76; p < 0.001$). The younger age groups are of course where we expect to find the elite or near-elite runners, who have the experience and fitness level to manage to keep a consistent pacing for a full half-marathon, but perhaps also many inexperienced recreational runners who start too fast and hit the wall.

### 3.2.3. Ability

We find an increasing share of runners the slower the finishing time for both men and women (see Supplementary Material Figure 7). For men, the share increases sharply for finishing times above 120 minutes, from less than 5% to over 30% among those finishing in over 150 minutes. The increase is less steep for women: for finishing times under 135 minutes less than 3% hit the wall, increasing to 25% for those finishing in over 180 minutes.

Conversely, the share of runners managing a negative split is highest among those finishing in 75-104 minutes for men (15-16%) and 90-119 minutes for women (14%), and then drops to 2-3% among the slowest runners. Note that among the very fastest group very few (<5%) run a negative split, possibly because of race tactics and placement being more important than finishing time.

### 3.3. Effect of Temperature

With higher temperature there is a trend towards both higher finishing times and a larger proportion of runners hitting the wall. The average finish time and proportion of runners hitting the wall is lower in the five coolest years studied (<18°C, small variation between years). As temperature increase, many runners manage to compensate for the by reducing their tempo (see Figure 3a, female: $r^2 = 0.90$; male: $r^2 = 0.91; p < 0.001$). The difference in average finishing time between the coldest (2012: 13.6°C) and the warmest (2013: 25°C) years is 7:36 minutes for women, and 9:18 minutes for men.

In warmer years we see an increased share of male runners hitting the wall ($r^2 =
0.85; \( p < 0.001 \), see Figure 3b. For women, the effect of temperature appears to be less of a factor \( (r^2 = 0.66, p = 0.004) \). Regarding negative splits, data appears to fall into two clusters representing the five cooler years \(<18^\circ\)\) and the five warmer years \(> 18^\circ\). Runners are about twice as likely (women: \( OR = 1.74 \); men: \( OR = 2.07 \)) to manage a negative split in the five cooler years (see Supplementary Material, Figure 8).

### 4. Discussion

Increased digitalisation and internet availability of results- and weather data allows for easier large-scale studies of pacing pattern of recreational races. In our work, we have demonstrated this methodology on ten years worth of results from Gothenburgh Half Marathon, one of the largest half-marathons in the world, which also has the ambition of being the safest. We identified runners at risk of overreach, ”hitting the wall” as well as runners at low risk, pacing more evenly.

Our results show that the majority of runners loose time on the second half of the race, having started fast, and gradually slowing down, similarly to the smaller studies by Nikolaidis et al. (2019); Čuk et al. (2019). Unlike these, we did however see evidence of an ”end spurt” where runners increased the pace between the 20 km mark and the goal. Women seem to pace more evenly than men of the same capacity, and men were twice as likely as women to overreach and hit the wall. Smyth (2021) found similar patterns in marathon, although the effect size was smaller than in our work \( (OR = 1.4 \text{ vs } OR = 2.0) \). We found a higher proportion of younger runners hitting the wall (age-group 17-29 years), and the lowest share among middle aged runners (40-49 years), which is consistent with previous smaller studies on the marathon and half-marathon distance (Deane et al., 2015; Smyth, 2021; Čuk et al., 2020, 2019). Carlström et al. (2019) found that younger runners at Gothenburg Half Marathon (below the average age) were also more likely to require ambulance assistance. Studies on half-marathons in South Africa has however found females over the age of 50 to have higher risk of medical complications (Schwabe et al., 2014b), and that older females are less likely to finish the races (Schwabe et al., 2014a). Our data only consisted of runners completing the race so we cannot say if the lower rate of hitting the wall among females could be because they instead abandon the race in such situations or not.

Younger and male runners are at both sides of the pacing spectrum: men are also
slightly more likely to run a negative or even split than women, and that younger runners do so to a larger extent than older runners. High temperature has a negative effect on both finishing times and the proportion of runners managing a negative split time. Men seem more adversely affected by increased temperature and both loose more time and increase the risk of hitting the wall more than women. Carlström et al. (2019) found increased incidence of cases needing ambulances in years (2010-2016) with temperatures above 17°C. Our work finds the same patterns mirrored in the much larger cohort of runners that hit the wall during the race with the same notable peaks in 2010 and 2013 (see Supplementary Material, Figure 9). This indicates that race organisers may find the number of runners "hitting the wall", as calculated from results data only, useful to estimate also need of ambulance assistance. In the five years with temperatures below 18°C, average finish times are faster, fewer hit the wall, and larger share of runners manages a negative or equal split pace. This lends additional support for the conclusion of Carlström et al. (2019) that low-risk temperatures for half-marathons range between 13 – 18°C.

4.1. Limitations

Our definitions do not include any personal metrics such as heart rate, reasons for slowdowns are unknown, and the threshold for slowdown could be tweaked. Still, we believe these definitions serve as a good enough proxy for revealing trends in pacing, also seen in other studies. With access to more fine-grained data e.g., GPS traces (Berndsen et al., 2020), HR monitors and training history, the models could be made more exact. However, this incurs a cost of more involved data collection, and a risk of skewing data towards more ambitious runners carrying appropriate devices and recording their training history.

Conclusions and Further Work

Gothenburg Half Marathon is one of the words largest and gather many recreational runners, however, there seems to be room for improvement in pacing. Working with a large dataset opens possibilities to apply machine learning techniques to the data, for example, predicting which runners risk hitting the wall before they do so. Initial experiments correctly identifies 76% of runners hitting the wall, but also misclassifies many that do not (Atterfors, Lamm, & Johansson, 2022 - in press). For more accurate results, additional details and personal data, e.g. heart rate, training history and GPS trace will be necessary. Until then, we expect our results to help race organisers and recreational runners to mitigate common risks, and perhaps help run a more enjoyable half-marathon next time.

Funding

The work was funded by a seed-grant from Chalmers Area of Advance Health Engineering.

Declaration of Interest

The authors report there are no competing interests to declare.
Ethics Statement

All data in this study is already publicly available online. Furthermore, it does not concern any sensitive personal information, as defined in GDPR §9. As such it is except from the requirement of formal Ethical Approval from the Swedish National Ethics Authority (oral communication with Chalmers Data Protection Officer David Screiber, 17 June 2022). Furthermore, it exempt from collecting written informed consent due to the public nature of the data, and large number of individuals. Further information (in Swedish) about the Swedish law: https://etikprovningsmyndigheten.se/for-forskare/vad-sager-lagen/.

Data Sharing

The full raw dataset of results from Gothenburg Half Marathon is publicly available online https://reg.goteborgsvarvet.se/sok/resultatlista.aspx. We worked with data from of the underlying database provided to us by the Gothenburg Half Marathon organisation, which excluded names of participants (individuals had just an identifying AthleteID-number), but otherwise containing the same data as the website. For GDPR reasons we have been asked to not re-publish this data elsewhere. Researchers interested in accessing this data can get it directly from Gothenburg Half Marathon’s website.

References


Ćuk, I., Nikolaidis, P., Markovic, S., & Knechtle, B. (2019). Age differences in pacing in endurance running: Comparison between marathon and half-marathon men and women. *Medicina (Kaunas), 55*(8).
Table 1. Summary of the data by year, number of runners, percentage of female runners, average finishing times and percentage of runners having hit the wall or run a negative split respectively. The warmest year was 2013 (25°C) and the coldest 2012 (13.6°C).

<table>
<thead>
<tr>
<th>Year</th>
<th>#Runners</th>
<th>%F</th>
<th>°C</th>
<th>Average Time</th>
<th>% HTW</th>
<th>% Neg Split</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>F</td>
<td>M ± F</td>
<td>M ± F</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>37 982</td>
<td>29.0</td>
<td>21.7</td>
<td>02:03:59 ± 00:19:35</td>
<td>02:15:45 ± 00:19:22</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>42 838</td>
<td>30.8</td>
<td>16.6</td>
<td>01:57:06 ± 00:18:27</td>
<td>02:09:59 ± 00:18:46</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>42 838</td>
<td>31.2</td>
<td>13.6</td>
<td>01:56:34 ± 00:19:01</td>
<td>02:09:10 ± 00:19:05</td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>44 919</td>
<td>33.0</td>
<td>25.0</td>
<td>02:05:22 ± 00:19:53</td>
<td>02:16:46 ± 00:20:00</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>47 187</td>
<td>34.6</td>
<td>18.9</td>
<td>01:59:38 ± 00:20:24</td>
<td>02:13:30 ± 00:20:18</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>46 207</td>
<td>34.8</td>
<td>14.7</td>
<td>01:57:43 ± 00:20:00</td>
<td>02:10:45 ± 00:19:44</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>44 972</td>
<td>34.8</td>
<td>15.1</td>
<td>01:57:38 ± 00:20:00</td>
<td>02:11:16 ± 00:19:47</td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>42 252</td>
<td>34.5</td>
<td>13.9</td>
<td>01:57:27 ± 00:19:43</td>
<td>02:10:49 ± 00:20:03</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>39 911</td>
<td>34.5</td>
<td>20.0</td>
<td>02:00:24 ± 00:21:17</td>
<td>02:14:40 ± 00:21:42</td>
<td></td>
</tr>
<tr>
<td>2019</td>
<td>33 134</td>
<td>34.0</td>
<td>19.4</td>
<td>01:59:58 ± 00:22:24</td>
<td>02:14:26 ± 00:22:05</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>423 496</td>
<td>33.2</td>
<td>17.9</td>
<td>01:59:28 ± 00:20:14</td>
<td>02:12:33 ± 00:20:15</td>
<td></td>
</tr>
</tbody>
</table>

11
### Table 2. Share of male and female runners hitting the wall or managing a negative or equal split by age group.

<table>
<thead>
<tr>
<th>Age</th>
<th>#Runners</th>
<th>% F</th>
<th>% HTW M</th>
<th>% HTW F</th>
<th>% Neg Split M</th>
<th>% Neg Split F</th>
</tr>
</thead>
<tbody>
<tr>
<td>17-29</td>
<td>89,031</td>
<td>44.1</td>
<td>13.0</td>
<td>7.0</td>
<td>14.7</td>
<td>12.3</td>
</tr>
<tr>
<td>30-39</td>
<td>125,484</td>
<td>32.9</td>
<td>10.2</td>
<td>4.7</td>
<td>11.6</td>
<td>10.1</td>
</tr>
<tr>
<td>40-49</td>
<td>124,275</td>
<td>31.3</td>
<td>9.9</td>
<td>4.2</td>
<td>9.0</td>
<td>7.7</td>
</tr>
<tr>
<td>50-59</td>
<td>62,261</td>
<td>26.5</td>
<td>9.3</td>
<td>5.5</td>
<td>6.3</td>
<td>4.7</td>
</tr>
<tr>
<td>60+</td>
<td>19,272</td>
<td>17.3</td>
<td>9.9</td>
<td>5.4</td>
<td>4.4</td>
<td>3.0</td>
</tr>
</tbody>
</table>