

# Efeitos associados da pandemia de COVID-19: uma comparação mundial do treinamento de corredores de longa distância em 2019 e 2020

L. A. Afonseca, M. Duarte, R. N. Watanabe

<sup>1</sup>Engenharia Biomédica, Universidade Federal do ABC  
São Bernardo do Campo, SP, Brazil, 09606-070

leo\_afa@hotmail.com, marcos.duarte@ufabc.edu.br, renato.watanabe@ufabc.edu.br

**Abstract.** *In this paper, we report how the training of long-distance runners was affected by the COVID-19 pandemic. Through web scraping, we extracted information related to the running activities recorded in all months of 2019 and 2020 from the Strava platform. We analyzed 10,703,690 activities of 36,412 athletes from all over the world. In 2020, compared to 2019, there was a 7% decrease in the volume of running training and a 7% decrease in the number of runners. We also observed large variations in these variables throughout 2020, reaching 35% less running volume in September 2020.*

**Resumo.** *Neste artigo, relatamos como o treinamento de corredores de longa distância foi afetado pela pandemia COVID-19. Por meio de web scraping, extraímos informações relacionadas à corrida registradas em todos os meses de 2019 e 2020 da plataforma Strava. Foram analisadas 10.703.690 atividades de 36.412 atletas de todo o mundo. Em 2020, relativamente a 2019, houve uma diminuição de 7% no volume de treinamento de corrida e de 7% do número de corredores. Também observamos grandes variações destas variáveis ao longo de 2020, chegando a 35% menos volume de corrida em setembro de 2020.*

## 1. Introdução

A doença Coronavírus 2019 (COVID-19) impactou profundamente nossas vidas em 2020; até março de 2021, aproximadamente 120 milhões de casos foram confirmados e 2,6 milhões de mortes atribuídas a ela (World Health Organization, 2021). Entre as principais medidas preventivas adotadas contra a propagação da COVID-19, o distanciamento social e o isolamento tem como potencial efeito colateral dificultar a prática de atividade física por alguém (obviamente, dificuldades além da própria pandemia em si), justamente quando é tão necessário buscar uma melhor qualidade de vida e que ser fisicamente ativo é fundamental para atingir este objetivo (Fragala et al., 2019). Enquanto a evolução da adoção de medidas de restrições ao longo do 2020 pelos governos mundiais pode ser monitorada por métricas disponíveis publicamente na internet (Hale et al., 2021), não é claro como a prática de atividade física foi afetada em 2020 e se foi afetada pela pandemia de COVID-19. Como o impacto da pandemia de COVID-19 sobre a prática de atividade física ainda é um tema muito recente, artigos científicos reportam dados para uma janela temporal incompleta para o ano 2020 (Stockwell et al., 2021). Para contornar essa limitação recorreremos aos dados públicos disponibilizados pela empresa *Strava* atuante no monitoramento de atividade física do usuário por meio de smartphones ou smartwatches

com acelerômetros e GPS. Para ter acesso ao grande volume de dados foram necessárias técnicas de raspagem de dados automatizadas (Saurkar, Pathare e Gode, 2018), ou seja, extração de informações não estruturadas com o intuito de criação de um banco de dados estruturado para análise.

No presente estudo, nosso objetivo foi aplicar técnicas de raspagem de dados e mostrar a aplicabilidade das informações através de uma investigação de como o treinamento de corrida no mundo ao longo do ano 2020 foi afetado pela pandemia de COVID-19 em relação à atividade em 2019. Nos concentramos nas atividades de corrida por ser das mais populares e acessíveis para as pessoas que buscam um estilo de vida saudável (Stamatakis e Chaudhury, 2008) e com a particularidade de ser mais comumente praticada em ambiente externo, o que potencialmente sugere que será mais afetada pelas medidas de restrição adotadas.

## 2. Métodos

Esta investigação foi feita a partir de registros públicos do aplicativo *Strava* (Strava Inc., 2020a). O estudo analisou dados anonimizados disponíveis publicamente na internet e, de acordo com as Resolução N° 466 e N° 510 do Conselho Nacional de Saúde, o Comitê de Ética em Pesquisa da Universidade Federal do ABC dispensou o projeto de avaliação pelo Comitê.

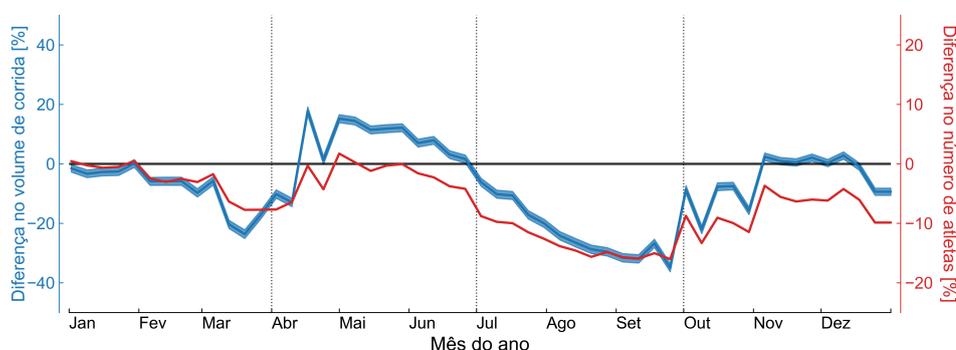
A extração dos dados foi realizada por meio de *web scraping*, técnica em que um software é escrito para implementar um robô web para automatizar o processo de acesso e extração de dados específicos de um site (Baskaran e Ramanujam, 2018). A raspagem dos dados é feita a partir do uso do operador *Selenium WebDriver* em conjunto com o navegador Google Chrome, que permite ações comuns da navegação web (Selenium, 2021), automatizada com a linguagem de programação Python. São pesquisados corredores de longa distância do *Strava* que tivessem ao menos um registro de ter corrido uma das seis maratonas Majors (Abbott World Marathon Majors, 2020) (as maratonas de Nova York, Londres, Boston, Tóquio, Chicago e Berlim) até 2019. Em seguida, todas as atividades dos usuários e informações relacionadas de janeiro de 2019 a dezembro de 2020 são extraídas. Esta delimitação foi necessária porque o *Strava* tem potencialmente milhões de usuários e seria inviável procurar em todas as páginas de usuários da plataforma para encontrar quem de fato é um corredor.

Na análise os dados são descritos com média e seu intervalo de confiança de 99% (CI99%) calculado entre os atletas em cada dia. O intervalo de confiança de 99 % foi calculado por *bootstrap* de percentil com 1.000 amostras aleatórias com repetição (Efron e Tibshirani, 1993).

## 3. Resultados

Foi realizada a análise dos registros de 10.703.690 atividades de corrida de 36.412 atletas de vários países em 2019 e 2020. Do total de atletas em 2019, 35083 correram pelo menos uma vez em 2020 (7% de redução). Os atletas correram em média cerca de uma vez a cada 2,5 dias e metade deles correu entre uma vez por semana e uma vez a cada dois dias em 2019 e 2020. Em 2019, o volume (a distância total percorrida) e duração total de corrida semanal por atleta foram em média 29 km e 161 minutos, respectivamente. Em 2020, esses números diminuíram para em média 27 km e 150 min, respectivamente (7%

de redução). Quanto aos atletas, 76% eram do sexo masculino; 34% tinham entre 18 e 34 anos, 59% entre 35 e 54 anos e os restantes 7% 55 anos ou mais. As regiões mais comuns de origem dos atletas eram: 38% dos Estados Unidos (de um total de 43% da América do Norte), 21% do Reino Unido (de um total de 44% da Europa), 8% da Ásia e 3% da América do Sul. A Figura 1 mostra que as diferenças relativas no volume de corrida semanal começaram na segunda metade do primeiro trimestre, onde se observa uma grande diminuição no volume de corrida em 2020 (em até 24% na terceira semana de março). Mas no segundo trimestre a diferença relativa foi no sentido oposto, o volume de corrida foi maior em 2020 neste período (em até 18% na terceira semana de abril). No terceiro trimestre de 2020 o volume de corrida volta a cair, desta vez de forma mais severa e por mais tempo, chegando a um mínimo de 35% menor na última semana de setembro. O comportamento da diferença relativa entre o número semanal de atletas correndo em 2019 e 2020 também foi qualitativamente similar ao padrão do volume de corrida.



**Figura 1. Comparação de 2020 em relação a 2019 referente ao volume de corrida (azul) e número de atletas (vermelho) versus data do ano. Dados normalizados pelo volume médio de corrida semanal em 2019 e número médio semanal de atletas em 2019 respectivamente.**

#### 4. Discussão

Considerando o ano inteiro, a prática de corrida de longa distância e o número de praticantes em 2020 não parecem ter sido muito afetados pela pandemia COVID-19, já que observamos uma redução média nestes números de apenas 7% de um ano para o outro. No entanto, houve grandes variações ao longo de 2020, chegando a 35% menos volume de corrida em setembro de 2020. Mas no segundo trimestre de 2020 observamos não só um retorno à prática de corrida, mas até um aumento delas em relação ao mesmo período de 2019. Isso talvez tenha acontecido como uma compensação pela diminuição do treinamento no mês anterior. Somente a partir do segundo semestre de 2020 é que a prática de atividade física volta a diminuir. Mas nos últimos meses de 2020 a prática de corrida volta a ficar mais parecida com 2019. As razões deste comportamento não são claras. Talvez as pessoas deixaram de seguir estritamente as políticas públicas ou as pessoas encontraram maneiras de praticar atividade física em acordo com as restrições, por exemplo, correndo sozinho, com máscara e em horários com menos pessoas nas ruas. Os resultados aqui reportados para o primeiro semestre de 2020 em relação a 2019, especificamente uma diminuição do volume de corrida no primeiro trimestre seguido por um aumento no

segundo trimestre, estão em acordo com outros relatos para atividades físicas em geral (Strava Inc., 2020b).

Nosso estudo contém limitações, o conjunto de dados analisado é uma amostra de conveniência não representativa de todos usuários do Strava. Considerando que as duas principais empresas no mercado de monitoramento e acompanhamento da prática de atividades físicas, *Fitbit* e *Strava*, relatam ter milhões de usuários cada uma, o *dataset* que utilizamos, com dados de cerca de 37 mil atletas, pode ser considerado pequeno demais. No entanto, este *dataset* é específico por conter apenas dados de pessoas que correram no mínimo uma das principais maratonas mundiais. Ainda que a diferença média no treinamento de corrida de longa distância entre os anos 2019 e 2020 possa ser considerada pequena, houve grandes variações ao longo de 2020, provavelmente devido à pandemia de COVID-19.

## Referências

- World Health Organization (20 de abr. de 2021). *Coronavirus disease (COVID-19) pandemic*. URL: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019> (acesso em 20/04/2021).
- Fragala, Maren S. et al. (2019). “Resistance training for older adults: Position statement from the national strength and conditioning association”. Em: *Journal of Strength and Conditioning Research* 33.8, pp. 2019–2052. ISSN: 15334295. DOI: 10.1519/jsc.0000000000003230.
- Hale, Thomas et al. (9 de fev. de 2021). *Oxford COVID-19 Government Response Tracker*. URL: <https://ourworldindata.org/covid-government-stringency-index> (acesso em 09/02/2021).
- Stockwell, Stephanie et al. (2021). “Changes in physical activity and sedentary behaviours from before to during the COVID-19 pandemic lockdown: a systematic review”. Em: *BMJ Open Sport & Exercise Medicine* 7.1. DOI: 10.1136/bmjsem-2020-000960. eprint: <https://bmjopensem.bmj.com/content/7/1/e000960.full.pdf>. URL: <https://bmjopensem.bmj.com/content/7/1/e000960>.
- Saurkar, Anand V., Kedar G. Pathare e Shweta A. Gode (2018). “An Overview on Web Scraping Techniques and Tools”. Em:
- Stamatakis, E e M Chaudhury (2008). “Temporal trends in adults’ sports participation patterns in England between 1997 and 2006: the Health Survey for England”. Em: *British Journal of Sports Medicine* 42.11, pp. 901–908. ISSN: 0306-3674. DOI: 10.1136/bjsem.2008.048082. eprint: <https://bjsm.bmj.com/content/42/11/901.full.pdf>. URL: <https://bjsm.bmj.com/content/42/11/901>.
- Strava Inc. (2020a). *Strava website*. URL: <https://www.strava.com/> (acesso em 10/12/2020).
- Baskaran, Umamageswari e Kalpana Ramanujam (2018). “Automated scraping of structured data records from health discussion forums using semantic analysis”. Em: *Informatics in Medicine Unlocked* 10, pp. 149–158. ISSN: 2352-9148. DOI: 10.1016/j.imu.2018.01.003. URL: <http://www.sciencedirect.com/science/article/pii/S2352914817302253>.
- Selenium (2021). *Selenium Webdriver*. URL: <https://www.selenium.dev/documentation/en/webdriver/> (acesso em 06/03/2021).
- Abbott World Marathon Majors (10 de dez. de 2020). *Abbott world marathon majors*. URL: <https://www.worldmarathonmajors.com/> (acesso em 23/06/2020).

Efron, Bradley e Robert J Tibshirani (1993). *An introduction to the bootstrap*. Chapman & Hall, p. 450.

Strava Inc. (2020b). *2020 Year In Sport*. URL: <https://blog.strava.com/press/yis2020/> (acesso em 16/12/2020).