The resumption of sports competitions after COVID-19 lockdown: The case of the Spanish football league.

Javier M. Buldú[a,b,c,d], Daniel R. Antequera[a,b,c], Jacobo Aguirre[b,e]

[a] *Complex Systems Group, Universidad Rey Juan Carlos, Madrid, Spain*
[b] *Grupo Interdisciplinar de Sistemas Complejos (GISC), Spain*
[c] *Laboratory of Biological Networks, Center for Biomedical Technology, Universidad Politécnica de Madrid, Madrid, Spain*
[d] *Institute of Unmanned System and Center for OPTical IMagery Analysis and Learning (OPTIMAL), Northwestern Polytechnical University, Xi'an 710072, China*
[e] *Centro de Astrobiología (CSIC-INTA), Ctra. de Ajalvir km 4, 28850, Torrejón de Ardoz, Madrid, Spain.*

Abstract

In this work we made use of a stochastic *SEIR Susceptible-Exposed-Infectious-Recovered* model and adapted it conveniently to describe the propagation of COVID-19 during a football tournament. Specifically, we were concerned about the re-start of the Spanish national football league (*"La Liga"*), which is currently stopped with 11 fixtures remaining. Our model included two additional states of an individual, namely, confined and quarantined, which can be reached when an individual has undergone a COVID-19 test with a positive result. The model also accounted for the interaction dynamics of players, considering three different sources of infection: (i) the player social circle, (ii) the interaction with his/her team colleagues during training sessions and (iii) the interaction with rivals during a match. First, we investigated how the time between matches influences the number of players that may be infected at the end of the season. Second, we simulated the effects of increasing the frequency of tests, whose sensitivity was also analyzed. Finally, we discussed different strategies to minimize the probability that COVID-19 propagates in case the season of "La Liga" were re-started after the current lockdown.

Keywords: Epidemics, SEIR, COVID-19, Sports

1. Introduction

The propagation of the virus SARS-CoV-2 officially started at the beginning of December 2019 in Wuhan (China), where the first COVID-19 victim was diagnosed with a new type of coronavirus. The virus first spread over different states in China before reaching other countries. On March 11th 2020, the World Health Organization (WHO) declared COVID-19 a pandemic, pointing to more than 118000 cases of the coronavirus illness in over 110 countries around the world [1]. The evolution of the pandemic, which is still affecting many countries worldwide, has been a matter of debate, since different strategies can be adopted to mitigate the spreading of COVID-19, some of them with unclear or unpredictable consequences. Due to the novelty of this unforeseen pandemic, the use of mathematical models is being extremely useful to predict the dynamics of the coronavirus spreading and the effects of different policies on the eventual reduction of the number of affected individuals.

Despite there are different approaches for modeling the pandemics, both continuous-time and discrete-time SIR-based models are probably the most extended approaches. The Suceptible-Infected-Recovered (SIR) model was first proposed by Kermack and McKendrick in 1927 [2] and consists of a compartmental model where individuals are split into three different states: (i) Susceptible (S), when they are sane, (ii) Infected (I), when they have the virus and (iii) Recovered (R). More sophisticated models include more posible states, such as Deceased in the SIRD model [3] or Exposed (E) in the SEIR model. The latter model has been extensively applied to describe the exponential growth of the number of individuals infected by SARS-CoV-2, the effects of quarantine and confinement measures and, ultimately, to evaluate the most adequate way of leaving confinement measures without increasing the risk of a second outbreak [4, 5, 6, 7, 8, 9, 10].

In this manuscript, we developed a discrete-time SEIR-type mathematical model that describes the spreading of the coronavirus during a sports competition. The motivation behind our study is that there has been a lively debate about how sports competitions that were not completed before the coronavirus crisis should be re-started or, ultimately, cancelled [11, 12, 13]. However, to the best of our knowledges, this debate has not being confronted with mathematical models that describe the propagation of SARS-CoV-2 between athletes. Here, we are concerned about the eventual re-start of the Spanish national league, which is currently suspended with 11 pending fixtures. We designed a mathematical model that incorporates the interaction of players during training sessions, leading to intra-club spreading, and during matches, responsible of inter-club contagions. Furthermore, we incorporated the use of tests to evaluate its consequences in identifying and confining those players that already have been infected. The model, whose main parameters were based on the scientific literature concerning the infection and recovery periods of COVID-19, could be easily adapted to describe other kind of sports competitions just by modifying the number of players and teams participating in the tournament.

2. Methodology

In SEIR models[14], a disease propagates through a network of individuals whose dynamical state can be either Susceptible S (healthy and susceptible to be infected), Exposed E (infected but in the latent period and therefore unable to infect other individuals), Infectious I (infected and able to infect other individuals), and Removed R (which includes (i) recovered individuals after having suffered the infection and therefore immune and (ii) deceased people).

Figure 1 represents a sketch of the model. The individuals (players from now on) can be infected at any time from people different to the players (technical staff of the team, family, etc.) with a probability β_{ext}. The second source of infection occurs during the training period, where they can be infected from other players of their own team with probability β_{train}. Finally, during the matches, players are exposed to infection from the players of their own team and the adversary team with probability β_{match}.

Once a player has been infected and becomes exposed, he/she has a probability σ of finishing the latent period and become infectious. Exposed and infectious players have, respectively, probabilities μ_E and μ_I of being detected as infected by COVID-19 via a virus test or because they show disease symptoms. If this is the case, players will be confined at their homes remaining in two possible states: exposed E_C or infectious I_c. Asymptomatic infectious players (belonging to class I, but not detected by virus tests), and confined infectious players, overcome the disease with probability γ. Note that confined players that have been recovered will remain quarantined (class Q) during a convalescence until they are prepared for playing again and become recovered (R) with probability γ_Q.

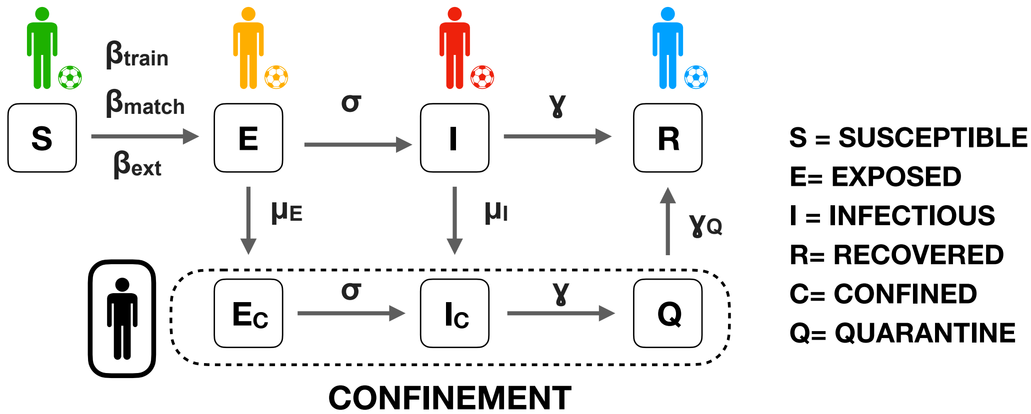

Figure 1: Schematic representation of the SEIR model adapted to a football competition. Players can be in different states: Susceptible (S), Exposed (E), Infectious (I) and Quarantined (Q). In case players are detected to be infected by the virus, they remain confined (indicated by the C suffix). After confinement, players will undergo a quarantine before being eligible to play again. Parameters β_{train}, β_{match} and β_{ext} account for the probability of becoming Exposed (E) during training, matches or externally (player social circle) respectively. Probability σ describes the transition from Exposed to Infectious (I). Probability γ controls the transition from Infectious to Recovered (R) or Quarantined (Q). Finally, γ_Q is related to the quarantine period a player must follow after recovery.

The days between virus tests N_{test} and the days between matches N_{match} are two critical variables for controlling the number of infected players during the championship, and therefore their influence in the model should be studied carefully. Note that the virus tests should be done in this context via Polymerase Chain Reaction (PCR) controls. The reason is that fast antibody or antigen detectors are only reliable more than a week after the infection, and in many cases even after the patient has already shown symptoms. This fact would allow the infectious (but not identified as infected) players to spread the virus during several days, making the control of the disease a hard task.

Modeling the Spanish national league

Our model can be applied to a diversity of competitions related to team contact sports, but we have focused on the re-start of the Spanish male national league. Therefore, we considered a competition with $M = 20$ teams composed of $L = 25$ players, the latter being the upper limit of players that can be registered by a team in the competition. The generalization to *Liga*

Iberdrola (Spanish feminine first division football ligue, with 16 teams), to the masculine or feminine football leagues of other countries, or even to other team sports (such as basketball, handball, rugby, etc.) is direct. Every team plays a match every N_{match} days, and during the $N_{match} - 1$ days in between the players train at their own stadiums. We supposed no resting days, as there is a clear interest for finishing the leagues as soon as possible, but including them in the model is trivial. We represented the training dynamics of the players, and contacts between them, using social networks instead of mean-field contacts. In this way, players' social networks during the team training followed a random structure of connections and were generated using an Erdös-Rényi model [15] with a probability $p = 0.2$ of connecting two players. This was done to describe the internal professional and friendship dynamics that every player has during training times and also during lunch time, etc. During training time, the infectious players (class I) might infect their neighbours in the social network with probability β_{train}. During the match day, every infected player on the pitch can infect any other player of its own team or the adversary with probability β_{match} (here we used a mean-field approach due to the inevitable contact dynamics that players follow during a match). Note that players cannot avoid voluntarily the contact with other players in the contest (with the exception, perhaps, of celebrating a goal, that could be forbidden if necessary), and therefore the contagion probability during a match might be larger than expected at first glance.

There is a wide range of values in the recent literature regarding each of the parameters that define the different steps of the disease (see Table 1 for a summary of the parameters of the model). The latent period σ^{-1} is the average time from infection to infectiousness, the incubation period is the average time from infection to the appearance of the first symptoms, and the infectious period γ is the average time that the patient is infectious. Depending on the virus, the latent period can be shorter or larger than the incubation period. In the case of COVID-19 the latent period is 1 or 2 days shorter in average than the incubation period, which makes it specially easy for the disease to spread among the population during the time in which people are infectious but asymptomatic. Regarding the mean incubation period, in [18] it was shown to be around 5 days, similar to that of SARS, and in [19] it was affirmed that it could be as short as 4 days. Note, however, that this quantity was not used in our model.

In [16] it was used a mean latent period σ^{-1} of 3 days and a mean infection period of $\gamma^{-1} = 5$ days, based on the Wuhan data. We selected these values

Parameter	Value (days^{-1})
β_{exp}	1/100000
β_{train}	1/10
β_{match}	1/100
σ	1/3; Refs. [16, 17]
γ	1/5; Refs. [16, 17]
γ_Q	1/5
μ_E	$\mu_I/3$
μ_I	[0,1], 0.9
N_{match}	[3-7]
N_{test}	[1-7]

Table 1: **Summary of the main parameters of the model.**

because they were also used in other more recent studies [17]. Note, however, that these are mean values: in [20] it was shown that the probability that patients with mild symptoms infected other people was very low after a week from the appearance of symptoms, but these means that in mild cases of COVID-19 patients can be infectious for as much as 10 days.

The values of the probability μ_I of being detected as infectious, either because a player shows disease symptoms or because the virus test yields a positive result, has been considered to be within the window [0,1], being 0 in case of not doing any test and being asymptomatic and 1 when tests have 100% sensitivity. However, when sensitivity of the test is not analyzed, we considered a value of 0.9 which is the typical one attributed to PCR tests. Concerning the probability of detecting an exposed individual, we set it as $\mu_E = \mu_I/3$, i.e. three times less than detecting an infected individual through the same test. The reason is that the viral load of an exposed individual is lower than that of an infectious one, therefore reducing the probability of a positive test result.

Furthermore, we have fixed the quarantine period σ_Q^{-1} to be 5 days, but varying slightly this quantity would not affect the model substantially. Also, there is no clear experimental data to fix the infecting probabilities β_{exp}, β_{train} and β_{match}. The probability β_{ext} of being infected during a day from people different to the players will slowly decrease as more and more people in the country recover from the disease, but for simplicity we have supposed it constant during the whole league. The probabilities of being infected during a training day or a match have been fixed at moderate values and we checked that slightly varying them did not qualitatively change the results.

Finally, we simulated between 10^3 and 10^5 seasons using our discrete-time model and obtained the main statistics of the accumulated number of infected players at time t, $n(t)$. Importantly, the seed of all simulations contained one player of the league who is already infected at the first day of the tournament (i.e., $n(0) = 1$). By doing so, the epidemic spreading begins at day one instead of any random day of the season, and therefore time t should be understood as days after the first infection.

3. Results

Figure 2 analyzes the influence that the number of days between tests and matches, N_{test} and N_{match}, have on the accumulated number of infected players $n(t)$ along the rest of the season (i.e. 11 matches and the training days in between). 1000 independent simulations were performed and the mean values of $n(t)$, $\bar{n}(t)$, are plotted in the figure.

In Fig. 2A we see how the mean accumulated number of infected players $\bar{n}(t)$ changes when the number of days between matches N_{match} is modified within the interval $\{3, 4, 5, 6, 7\}$, i.e., we set the minimum and maximum number of days between matches to 3 and 7, respectively. Interestingly, we can see that it is convenient to reduce the time between matches to the minimum. The reason is twofold. On the one hand, with N_{match} being the lowest, the competition would last fewer weeks and therefore the players would be exposed for less time. On the other hand, the probability of being infected is higher during a training day than during a match day, since players are more exposed to physical contact with other players during trainings. For this reason, the higher the number of days between matches, the higher the slope of the curves of Fig.2A.

In Fig. 2B we show the different evolution of the mean value of the accumulated number of infected players $\bar{n}(t)$ when PCR tests are or are not performed. Matches are played with a separation of 7 days, in this case. We can observe how skipping the tests increases substantially the number of infected players, whose growth is specially higher at the beginning of the spreading process. These results show that conducting a coronavirus detection test is essential to prevent its spread among "La Liga" teams.

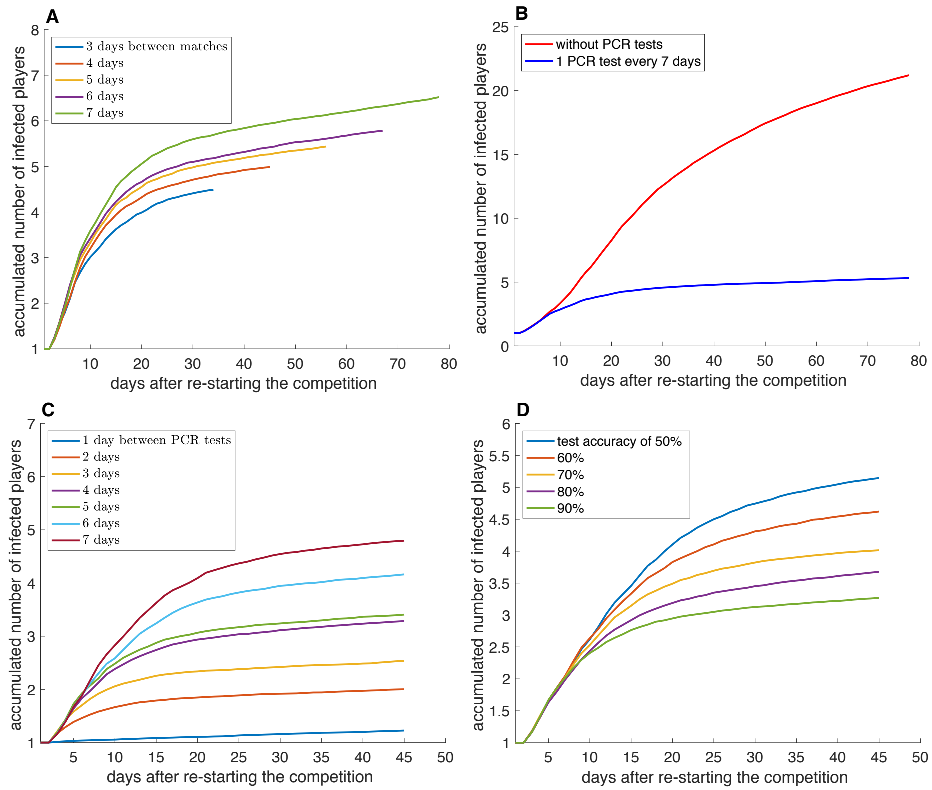

Figure 2: Dependence of the mean of the accumulated number of players that have been infected at the end of the season $\bar{n}(t)$ on the main parameters of the system: the days between matches (N_{match}) (A) and the days between PCR tests (N_{test}) (B-D). We simulated 1000 times the rest of the season, that consists of 11 matches and the training days in between. Parameters of the simulation are indicated in Table 1, with $\mu_I = 0.9$, unless specified otherwise. The seed of all simulations contained one player infected at the first day of the tournament. (A) Influence of the number of days between two consecutive matches, N_{match}, on $\bar{n}(t)$. In this simulation, PCR tests with 90% sensitivity were carried out every $N_{match} = 7$ days. In (B) we compare the outcome of not doing any tests during the rest of the season and doing them every $N_{test} = 7$ days (matches played every 7 days), while in (C) we focus on the number of days N_{test} between each PCR control (matches played every 4 days, closer to the optimum frequency of 1 every 3 days). (D) Influence of test accuracy μ_I on $\bar{n}(t)$ (PCR tests and matches carried out every 4 days).

However, it is necessary to take into account the frequency and reliability of such tests. To investigate this issue, we assume that it is decided to play, for example, every 4 days, a measure close to the most favorable scenario of 3 days, although not so extreme. In Fig. 2C we see how important it is to perform tests as often as possible, ideally every day. As the tests are more separated over time, the risk of infecting more players inevitably increases. Finally, it is possible to simulate how important the accuracy of the tests is and the consequences of making use of low sensitive methods. Figure 2D shows how the average number of infected players $\bar{n}(t)$ increases as the reliability of the tests μ_E and μ_I decrease. These results support the convenience of performing PCR testing, whose accuracy is estimated to be close to 90% and larger than any other method.

As mentioned below, the curves shown in Fig. 2 are the mean values $\bar{n}(t)$ obtained after 1000 simulations of the model. While the standard deviation of the mean $\bar{n}(t)$, $\sigma_{\bar{n}}$, is too small to be distinguished in any of the plots, the standard deviation of $n(t)$, σ_n, is on the contrary very large –in some cases of the order of the mean $\bar{n}$– and shows that the evolution of a single process is highly unpredictable. To cast some light on this question, in Fig. 3 we have plotted the probability function of the accumulated number of infected players $n(t)$ (i.e., probability of obtaining $n(t) = 1, 2, 3...$ accumulated infected players after t days, calculated as the normalized histogram of 10^5 simulations of the process) when matches and PCR controls are carried out every 7 days (green curve in Fig. 2A) after $t = 4$ days, 20 days, and at the end of the league ($t = 78$ days). In the first days of the competition, the disease starts to spread in the team of the so-far unique infected player. As expected, a Poisson distribution approximates accurately its function probability. However, the disease soon spreads towards other teams and the function distribution becomes more complex: at moderate values of the time ($t = 20$) the probability function presents a *hump* that certifies that the curve is in fact the consequence of several spreading processes, that is, the addition of intra- and interteam spreading, plus the potential infections coming from outside the league. Finally, when the season reaches its end ($t = 78$) the curve already presents several humps, as well as an exponential-like tail, and in fact at that time the standard deviation σ_n is almost as large as the mean $\bar{n}$ (as it is verified in exponential probability density functions). Note that, while a normal approximation is not accurate at this time, when t grows substantially (many weeks after the end of the season, not shown) the probability function becomes a Gaussian, as expected from the Central Limit Theorem.

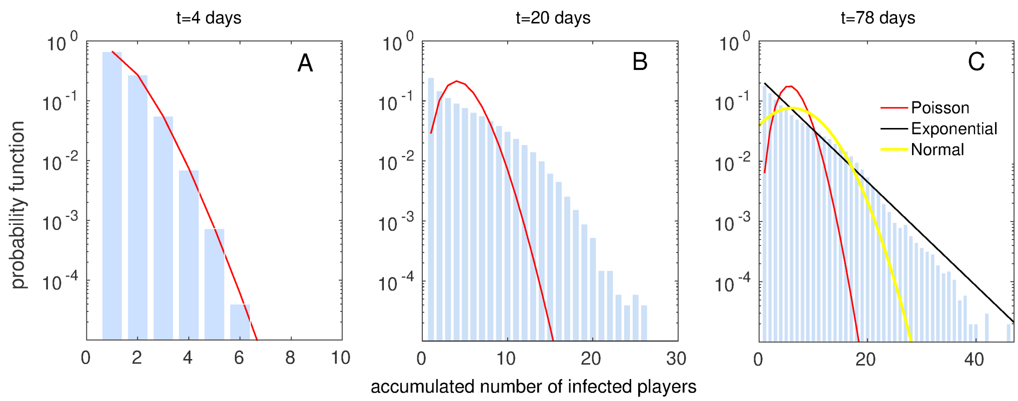

Figure 3: Probability function of the accumulated number of infected players $n(t)$ at the beginning of the re-start (A, $t = 4$ days), after few weeks (B, $t = 20$ days) and at the end of the season (C, $t = 78$ days), when matches and PCR controls are carried out every 7 days. For these 3 cases, $\bar{n} \pm \sigma_n = 1.4 \pm 0.6$, $\bar{n} \pm \sigma_n = 4.5 \pm 3.6$, $\bar{n} \pm \sigma_n = 6.0 \pm 5.2$, respectively. Approximations to a Poisson distribution of mean $\bar{n}$, to an exponential distribution of mean $\bar{n}$ and to a normal distribution of mean $\bar{n}$ and standard deviation σ_n are shown. 10^5 simulations of the league were performed.

4. Conclusions

"All the models are wrong, but some are useful". This famous statement, attributed to the statistician George Box, sums up the usefulness of our model. Although it is not possible to predict the number of infected individuals, the model allows to describe in a qualitative way the influence that different measures can have to mitigate the spreading of the coronavirus during a competition. Based on the simulations carried out with the epidemiological parameters estimated by the scientific community, the results of the study can be summarized in five points:

- Reducing the days between matches reduces the risk of spreading COVID-19 throughout the season. The more the season is compressed, the less risk of contagion.

- PCR tests should be performed on all football players participating the competition. Antibody and antigen tests should be ruled out in this context because they are less reliable and are not effective until the disease is well advanced.

- The tests should be carried out continuously along the competition, with the optimum scenario being one test per day.

- The player's environment is essential to avoid introducing the disease into the system. It is necessary that the players try to limit their social contacts as much as possible, and that their physical interaction with the technical staff is as distant as possible.

- The process is highly unpredictable. While qualitative results are clear, obtaining precise predictions for a single realization (the real case) is not possible.

We must also note that applying all the measures suggested by the model involves a cost. On the one hand, reducing the time between matches can be a problem for players from a physical point of view. The physical recovery time after a match would be reduced and the risk of injury would increase. To reduce this risk, teams should increase player rotations. Regarding the tests, football clubs should provide the necessary support and means to carry out such a high number of tests in such a short time. Without an efficient policy in this regard, the risk of reinfection in competition would skyrocket. Players will also pay a personal cost to control the eventual spreading of coronavirus. Minimizing their contacts with other individuals would mean limiting their travels, public events and, in general, reducing interactions with people outside the family environment. In fact, maintaining them in a confinement during the rest of the season would be obviously the optimum situation.

Finally, although the results shown here are focused on the resumption of the men's Spanish national league, the conclusions are equally valid for the women's competition. Furthermore, the model could be adapted to any competition in which matches involves some physical contact between players, such as basketball, handball or rugby.

Acknowledgements

The authors acknowledge the assistance of E. Lázaro and J.A. Sánchez-Brunete. JMB is funded by MINECO, Spain (FIS2017-84151-P). DRA is supported by Comunidad de Madrid through project MPEJ-2019-AI/TIC-13118. JA is supported by MINECO, Spain (FIS2017-89773-P, MiMevo), and the Spanish State Research Agency (AEI) Project No. MDM-2017-0737

Unidad de Excelencia “María de Maeztu”-Centro de Astrobiología (INTA-CSIC).

References

[1] https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020.

[2] Kermack WO, McKendrick AG, A contribution to the mathematical theory of epidemics. Proc. R. Soc. London 1927; 115:700.

[3] Matadi MB. The SIRD epidemial model. Far East Journal of Applied Mathematics 2014; 89;1-14.

[4] Peng L, Yang W, Zhang D, Zhuge C, Hong L. Epidemic analysis of COVID-19 in China by dynamical modeling. arXiv preprint arXiv:2002.06563.

[5] Radulescu A, Cavanagh K. Management strategies in a SEIR model of COVID 19 community spread. arXiv preprint arXiv:2003.11150.

[6] Hou C, Chen J, Zhou Y, Hua L, Yuan J, He S, Guo Y, Zhang S, Jia Q, Zhao C, Zhang J, Xu G, Jia E. The effectiveness of the quarantine of Wuhan city against the Corona Virus Disease 2019 (COVID-19): well-mixed SEIR model analysis. J Med Virol 2020.

[7] Rovetta A, Bhagavathula, AS. Modelling the epidemiological trend and behavior of COVID-19 in Italy. medRxiv 2020.03.19.20038968; doi: https://doi.org/10.1101/2020.03.19.20038968

[8] De Brouwer E, Raimondi D, Moreau Y. Modeling the COVID-19 outbreaks and the effectiveness of the containment measures adopted across countries. medRxiv 2020.04.02.20046375; doi: https://doi.org/10.1101/2020.04.02.20046375

[9] Ngonghala CN, Iboi E, Eikenberry S, Scotch M, MacIntyre CR, Bonds MH, Gumel AB. Mathematical assessment of the impact of non-pharmaceutical interventions on curtailing the 2019 novel Coronavirus. medRxiv 2020.04.15.20066480; doi: https://doi.org/10.1101/2020.04.15.20066480

[10] Lopez L, Rodó X. The end of the social confinement in Spain and the COVID-19 re-emergence risk. medRxiv 2020.04.14.20064766; doi: https://doi.org/10.1101/2020.04.14.20064766

[11] Halabchi F, Ahmadinejad Z, Selk-Ghaffari M. COVID-19 Epidemic: Exercise or Not to Exercise; That is the Question! Asian J Sports Med. 2020;1(1):e102630.

[12] Duarte Muñoz M, Meyer T. Infectious diseases and football - lessons not only from COVID-19. Science and Medicine in Football 2020;4:85-86.

[13] Corsini A, Bisciotti GN, Eirale C, Volpi, P. Football cannot restart soon during the COVID-19 emergency! A critical perspective from the Italian experience and a call for action. Br J Sports Med Epub ahead of print: [28/04/2020].

[14] Newman M, Networks: An Introduction. Oxford University Press, Inc.; 2010.

[15] Erdös P, Rényi A. On random graphs I. Publicationes mathematicae 1959; 6(26):290–297.

[16] Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. The Lancet 2020; 395(10225):689–697.

[17] Lin Q, Zhao S, Gao D, Lou Y, Yang S, Musa S *et al.* A conceptual model for the coronavirus disease 2019 (COVID-19) outbreak in Wuhan, China with individual reaction and governmental action. International Journal of Infectious Diseases 2020; 93(26):201–216.

[18] Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR *et al.* The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: Estimation and application. Ann Intern Med 2020; doi: https://doi.org/10.7326/M20-0504.

[19] Guan W, Ni Z, Hu Y, Liang W and Ou C, He J *et al.* Clinical Characteristics of Coronavirus Disease 2019 in China. N Engl J Med 2020; DOI: 10.1056/NEJMoa2002032.

[20] Wölfel R, Corman VM, Guggemos W, Seilmaier M, Zange S. Müller M *et al.* Virological assessment of hospitalized patients with COVID-2019. Nature 2020; https://doi.org/10.1038/s41586-020-2196-x.