Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Spatial risk assessment of an emerging pandemic under data scarcity: A case of COVID-19 in Eswatini

Wisdom M. Dlamini a, Sabelo N. Dlamini a,⇑, Sizwe D. Mabaso a, Sabelo P. Simelane b

a University of Eswatini, Department of Geography, Kwaluseni, Manzini, Eswatini
b Central Statistics Office, Ministry of Economic Planning, Mbabane, Eswatini

ARTICLE INFO

Keywords: Eswatini Transmission risk Susceptibility risk COVID-19 Exposure risk

ABSTRACT

Coronavirus (COVID-19) has rapidly spread across many countries in pandemic proportions since the first case was reported in Hubei, China in December 2019. Understanding transmission, susceptibility and exposure risks is crucial for surveillance, control and response to the disease. Knowing the geographic distribution of health resource scarcity areas is necessary if a country is to adequately anticipate and prepare for the full impact of infections. We explored the potential to undertake a spatial risk assessment of an emerging pandemic under data scarcity in Eswatini. We used a set of socio-economic and demographic variables to identify epidemic risk prone areas in the country. Three risk zone levels for COVID-19 were identified in the country. The analysis showed that about 29% (320 818) of the population were located in the high risk zone and these were people who could potentially be infected with COVID-19 in the absence of mitigation measures. A majority of cases and deaths attributed to COVID-19 would likely remain unknown but our estimate could be used to gauge the full burden of the disease. Approximating and quantifying the number of people who may be potentially infected with COVID-19 remains impossible under data scarcity and limited healthcare capacity especially in sub-Saharan Africa. We provided an estimation method that could support the pandemic risk forecasting, preparedness and response measures in the midst of data scarcity. The resultant map products could be used to guide on-the-ground surveillance and response efforts.

1. Introduction

Coronavirus sickness (also known as COVID-19) is caused by a virus, more specifically, a coronavirus SARS-CoV-2, and has resulted in an outbreak of respiratory illness globally (Wu, Chen, & Chan, 2020). The virus has rapidly spread across many countries in pandemic proportions since the first case was reported in Hubei, China in December 2019 (Zhang et al., 2020). This led to the World Health Organization (WHO) declaring it a global pandemic in March 2020 (Kazory et al., 2020). The disease manifested itself differently within the affected populations, whereby it may also be fatal in its acute respiratory form. Understanding the transmission, susceptibility and exposure risks of an affected country is crucial in estimating the severity of impact in the affected populations as well as in planning preparedness and response strategies (Bedford et al., 2020). Knowing the geographic distribution of health resource scarcity areas is necessary if a country is to adequately anticipate and prepare for the full impact of this respiratory infection that has globally shattered many countries’ health systems, leading to untold humanitarian devastations (Kazory et al., 2020).

In Eswatini, the first case of COVID-19 was reported on 13th March 2020 and since then sporadic episodes of the disease have been reported in different areas around the country. Eswatini is a southern African country located in the north eastern part of South Africa and very close (about 90 km) to the southern part of Mozambique. The country is landlocked and it is bordered by South Africa all around except for the north-eastern side which is bordered by Mozambique. Eswatini is classified as a developing country and is 70% rural. Most rural folks derive their livelihoods from substance agriculture. The remaining 30% of the population resides in urban towns and cities and is mostly sustained by formal employment and small to medium formal and informal businesses. The urban areas of the country are characterized by high population densities compared to the rural parts which are not only sparsely populated but also have high poverty incidence.

Although the disease could infect anyone within the demographic structures of the population, its clinical symptoms vary tremendously from person to person (Wilder-Smith et al., 2020). Kazory et al. (2020)
national health system capacity, the study also mapped insufficient resource risk to highlight the limitations of active case detection in the country. The EAs are administratively used to plan and conduct national health in
susceptibility risk, transmission risk and exposure risk in Eswatini. The data were aggregated at enumeration area (EA) level which is estimated the potential number of people who may already be infected with COVID-19, albeit not officially reported in the country’s health information system database. 2. Methods 2.1. Data sources

Using spatial data for Eswatini and information of reported cases, modelling of COVID-19 transmission risk, susceptibility risk, insufficient resource risk, and exposure risk was undertaken. The datasets used for estimation of the four risks are listed in Table 1. The Exposure risk used the reported COVID-19 cases only. The demographic data were derived from the 2017 Population and Housing Census data obtained from the Central Statistics Office (CSO) (Central Statistics Office, 2018a) (see Table 2).

The data were aggregated at enumeration area (EA) level which is the smallest geographic unit used for census data collection in the country. The EAs are administratively used to plan and conduct national surveys and censuses and they range from an area of about 0.013 km² to about 194 km². The average population is about 470 people per EA with an average of 114 households per EA. Each household had an average of about 5 people and the population density was as low as 0.02 people per

| Table 1 | Datasets used in the risk analysis. |
|----------------------------------|-----------------------------------|
| Dataset                          | Definition                        | Description                                                                 | Source                      | Transmission risk | Susceptibility risk | Health resource scarcity risk |
| Human population density         | Number of people/per unit area     | Numerical quantities of the populated surface area in each EA.                | CSO (2018a)                 | ✓                | ✓                  |                         |
| Housing density                  | Number of buildings per unit area  | Numerical quantities of the built up surface area in each EA                 | Science Information Network - CEESIN - Columbia University, 2016 | ✓                |                     |                         |
| Annual average traffic density   | Number of vehicles moving through  | Numerical quantities of average traffic moving through each EA approximated as | Ministry of Public Works and Transport – 2017 data | ✓                |                     |                         |
| Population of elderly (55+ per 1000) | people                        | percentage or rate of people above 55 years of age in each EA                | CSO (2018a)                 | ✓                | ✓                  |                         |
| HIV prevalence                   | Percentage of population which is | Percentage of people living with HIV in each EA                             | SHIMS2                      | ✓                | ✓                  |                         |
| Poverty incidence                | Percentage of households living   | Percentage of people living below USD 2 per day in each EA.                  | CSO, 2011                   | ✓                | ✓                  |                         |
| Proximity to health facilities   | Distance to the nearest health    | Continuous numerical distance to health facility for each EA                 | Ministry of Health - 2020 data | ✓                | ✓                  |                         |
| Number of hospital beds          | Number of beds in the major referral hospitals and clinics | Numerical quantity of the total hospital beds available for patient occupation in each health facility | Ministry of Health - 2020 data | ✓                | ✓                  |                         |
| Employment rate                  | Percentage of adult population    | Rate of employment for the population located in each EA                     | CSO (2018a)                 | ✓                | ✓                  |                         |
| Population size                  | Number of people                  | Number of people in the entire country obtained by summing up the number of people recorded in each EA | CSO (2018a)                 | ✓                | ✓                  |                         |

Further noted that uncertainties surrounding various aspects of its optimal management strategies persist. Moreover, predicting the severity of the disease in terms of the number of the people likely to be impacted remains a challenge, especially when using standard models which rely on the absolute number and location of reported cases. This is because it had been shown that there is a significant proportion of people who recover from the disease without manifesting any symptoms (Yu and Yang, n.d.). Therefore, predictions of the extent of infections in space and time remains elusive and the actual burden of the disease may never be truly known.

Eswatini’s reported cases are picked via random screenings that are conducted on the roadside, and in rare occurrences, when an infected person present at a health facility. Once a case is confirmed contact tracing then ensues to ensure that possible cases are also proactively identified. Official data on COVID-19 cases is likely to be unreliable as it depends on the number of tests conducted per capita. There is also a high possibility that mild cases will self-treat and not be officially captured in the statistics of confirmed cases. Based on the official data on COVID-19 from the Ministry of Health, a higher percentage of confirmed cases are either asymptomatic or mild, and to date, there have been 45% recoveries in patients (Ministry of Health, 2020). This may be a strong indication that a majority of the infected will not present at a health facility, and may also be missed in the routine random screenings due to the ongoing partial lockdown measures implemented by the government as means to restrict movement.

This paper, therefore, sought to explore the potential to undertake a spatial risk assessment of an emerging pandemic under data scarcity. It used a set of socio-economic and demographic variables in order to identify epidemic risk prone areas in the country, as well as to map susceptibility risk, transmission risk and exposure risk in Eswatini (Townsend, 2015; J. T. Wu, Chen, & Chan, 2020). In view of the fact that contact tracing, testing, quarantine and isolation of cases depends on the national health system capacity, the study also mapped insufficient resource risk to highlight the limitations of active case detection in the country (Nkengasong & Mankoula, 2020). Lastly, this work then
km² and as high as 38 949 km². The country is made up of a total of 2326 EAs which are demarcated by boundaries such as rivers, roads, valleys, mountains and other similar natural and human-made boundaries.

The traffic volume (expressed as the number of vehicles passing through a specific point on the road network per day) was converted to traffic density by dividing the annual average daily traffic density (AADT) values with the land area of the enumeration polygon through which it passes. The traffic density was then used to quantify spatial interaction between EAs as a proxy for human contact. In the absence of other explicit human mobility data (such as mobile phone data), traffic density was found to be a better alternative measure of human mobility. We hypothesized that road traffic density could serve as a reflection of societal activity and, to an extent, the likelihood of personal interaction as similarly done in other studies (e.g. (Parr et al., 2020). This is also useful in the Eswatini where a majority of the population uses public transport whose usage has been observed to strongly correlate with daily and cumulative number of COVID-19 cases (Zheng et al., 2020). Housing density was estimated by calculating the land area with buildings using Facebook’s 30-m population product (Facebook Connectivity Lab and Center for International Earth Science Information Network - CIESIN - Columbia University, 2016). The dataset was binarized into building and non-building pixels using conditional if/else evaluation where pixels with zero estimated population were reclassified as having no buildings and vice versa.

HIV prevalence data was geostatistically interpolated from facility level test data. The Empirical Bayesian kriging (EBK) technique, which automatically computes interpolation parameters using a process of subsetting and simulations, was utilized for this task. The EBK approach has the advantage of having standard errors of predictions that are more accurate than other kriging methods. In addition it has a relatively more accurate prediction of ordinary nonstationary data and smaller datasets (Pilz & Spock, 2008). The EBK approach also considers the error introduced by approximating the basic semivariogram, in this case, the power semivariogram.

All the variables used in the analysis were first standardized using a z-transform where the mean for all values was subtracted from each value and divided by the standard deviation for all values in each variable. The standardization process allows for comparison of values from different distributions such as ratios, distances and population (ESRI, 2020). For each EA, the standardized values were subtracted from those of a worst-case target value and the differences were then squared and added together. This was done by comparing or ranking the variable values in each EA against a worst-case target polygon. This polygon contained the worst-case values for all the variables used for each risk variable used. This was created by randomly selecting one EA and copying it to a new feature class (polygon). Each variable in this target polygon was then edited to reflect the worst-case (highest risk) values found throughout the study area. For example, this polygon had the highest number of elderly people for any of the EAs, the lowest value for healthcare resources, the highest population density, the largest traffic density, and so forth.

The sum became the similarity index for that EA. After processing all the EAs, they were then ranked from smallest index (most similar) to largest index (most dissimilar). Ultimately, all the EAs were ranked based on the values of the target polygon from 1 (very similar the worst-case feature) to 2036 (very dissimilar to the worst-case feature). These values were converted to a range of 0–1 and reversed such that a value of 1 corresponds to the highest risk and 0 for the lowest risk.

3. Risk analysis

3.1. Transmission risk mapping

The most severe risk for COVID-19 transmission is often in locations associated with dense populations and high human contact and interaction (Rader et al., 2020; World Health Organization, 2020). In this case transmission risk is the likelihood that a contact between an infected person and a susceptible individual will result to the susceptible person being infected. We developed a transmission risk map by utilizing human population density, building density and a spatial interaction index derived from AADT (i.e. traffic volumes per square kilometre). We extracted demographic variables from the 2017 Population and Housing Census (Central Statistics Office, 2018a). The initial assumption was that areas with high transmission risk would normally be densely populated urban and peri-urban areas that would also be characterized by high traffic volumes and human interaction.

3.2. Exposure risk mapping

Mapping the severity risk of exposure to COVID-19 is essential in the fight against the disease. Exposure risk measures the potential exposure to an individual infected with disease in the country. We used the locality information for cases reported until the April 15, 2020 to geolocate the cases in order to develop the exposure risk map. Mapping of the cases was based on the reported localities using the EA names and approximated using local points of interest data. The 2326 EAs were the lowest-level appropriate unit of analysis ranging in size from 0.012 km² to a maximum of 194.19 km², averaging at 7.5 km². The large EAs were mostly large areas of sparsely populated plantation forests, sugarcane plantations, ranches and protected areas in the country. The locations of the cases were then systematically put within each EA and simultaneously as close as possible to buildings by using the building footprint data. This was cross-checked with Google Earth imagery to identify buildings and high-density populated areas within each EA using the locality name of the reported case in order to avoid misplacing the cases too far from populated areas. The identified buildings and high-density populated areas of each locality within an EA was used to approximate the geographic location of each case. We then estimated the proximity to the nearest 5% of the rest of the COVID-19 cases using the map of the geocoded cases.

3.3. Susceptibility risk mapping

Most data on COVID-19 from other countries had shown that many people who contract COVID-19 would have mild symptoms, and most would recover (Dong et al., 2020; Sun et al., 2020). Often children and younger adults have been found to generally recover quicker compared to older adults and those living with chronic illnesses who seem to markedly contribute to the observed mortality rates (Li et al., 2020). In addition, the risk seem to increase where large numbers of susceptible people live such as in densely populated areas and among those with limited resources to cater for their health and well-being. In this study susceptibility risk was defined as the likelihood of underlying health condition that could aggravate the severity of COVID-19 outcome on the affected population. We estimated susceptibility in this analysis by utilizing the population of elderly people (55+ years) per 1000, population density, estimated HIV prevalence, and poverty incidence. Ideally, we would have additionally used data on co-morbidities of the affected population. However, this data was not immediately available for this study during the data analysis. Only the HIV prevalence data was available at national level and it was used as an indicator for susceptibility in the present study. Nevertheless, other disease comorbidities are high on people living with HIV in Eswatini (Masuku et al., 2019; Rabbkin et al., 2018). The poverty incidence was used as a measure of poverty which is the proportion of the population that is below the poverty line of less than USD 2 per day as derived from the 2016/2017 Eswatini Household Income and Expenditure Survey (Central Statistics Office, 2018b).
3.4. Resource (healthcare) scarcity risk mapping

Healthcare resource scarcity is usually a major concern in public health and it is believed that the COVID-19 outbreak would strain healthcare resources, potentially intensifying the pandemic’s negative effects. Health resource scarcity in this case was the measure of resource availability and adequacy to enable a well-functioning health facility. Therefore, there is a need to identify which areas would be at highest risk of suffering from healthcare resource scarcity in the rest of the country. We estimated access to healthcare facilities by utilizing a gravity model which was based on the number of available beds as a measure of attractiveness. The gravity model approximated the probabilistic attraction an origin (EAs) would have towards a destination (health facilities) based on the distance between that origin and destination and the attractiveness (i.e. mass) of the destination. We used the Huff gravity model which takes the following form (Huff, 2003):

$$ P_{ij} = \frac{W_i}{\sum_{i} \left( \frac{W_i}{D_{ij}} \right)} $$

where:

- $P_{ij}$ = the probability of population in area $j$ using health facility $i$.
- $W_i$ = a measure of the attractiveness of each health facility $i$ (in our case, the number of beds).
- $D_{ij}$ = the distance from area $j$ to health facility $i$.
- $a$ = an exponent applied to distance so the probability of distant sites is dampened. A value of 2 was used in our model.

3.5. Risk profiling and clustering

In order to develop targeted interventions, information combining all the aforementioned risks is required. This could be a map showing areas that have similar risk profiles regarding COVID-19. Subsequently, a multivariate clustering was conducted to partition the country into zones with similar characteristics or risk profiles using the four similarity indices as inputs. We used a multivariate clustering approach that utilizes unsupervised machine learning methods to search for natural clusters in the data. The clusters were created using the k-means algorithm (Lloyd, 1982). The k-means algorithm segregates features so that the disparities within the features in a cluster, over all clusters, are reduced. It has a fast convergence with the added advantage of being mathematically uncomplicated, and straightforward to implement (Yuan & Yang, 2019). A greedier heuristic approach is used to cluster features in order to take care of the NP-hard algorithm. When determining clusters in the data, an $R^2$ value was computed for each risk category. This value reflects how much of the variation in the data was explained after the clustering process. The $R^2$ was computed as follows (ESRI, 2020):

![COVID-19 transmission risk map for Eswatini.](image-url)
We allocated a seed value of 1 to the EAs with the highest transmission risk, resource risk, susceptibility risk, and exposure risk, while all the remaining EAs were assigned a value of 0. The seed value represents a number used to initialize a random number generator and specifies a stream from a set of possible random numbers. This clustering approach searches for a mix where all the features within each cluster would be as similar as possible (i.e. having as similar indices as possible), and the remaining clusters that would be as dissimilar as possible. This seeding approach starts the search for an optimal result with the risk extremes and is often efficient as it ensures that the same result is consistent and obtained each time the analysis is undertaken.

\[ R^2 = \frac{\sum_{i=1}^{n_c} \sum_{j=1}^{n} \sum_{k=1}^{n_v} (V_{ikj} - \bar{V}_k)^2}{\sum_{i=1}^{n_c} \sum_{j=1}^{n} \sum_{k=1}^{n_v} (V_{ijk} - \bar{V}_k)^2} \]

\( n \) = number of features  
\( n_i \) = the number of features in cluster \( i \)  
\( n_c \) = the number of classes (clusters)  
\( n_v \) = the number of variables used to cluster features  
\( V_{ikj} \) = the number of the \( k \)th variable of the \( j \)th feature in the \( i \)th cluster  
\( \bar{V}_k \) = the mean value of the \( k \)th variable  
\( \bar{V}_k \) = the mean value of the \( k \)th variable in cluster \( i \)

4. Results and discussion

4.1. Risk mapping

The derived risk maps indicate areas with high (0) to low similarity (1) with the worst-case values estimated from the similarity indices. For instance, areas with high transmission risk are very similar in relation to the factors influencing transmission risk, i.e. they have high human population density, high housing density, and high annual average traffic density. The results showed that areas with high transmission risk were the densely populated urban and peri-urban locations where there was high spatial interaction as evidenced by high traffic volumes (Fig. 1). These also included areas that were closer to the main transit routes, that is between the country’s urban areas such as the major cities and towns. Clearly the risk of spread was expected to be more severe in this place due to high human interaction as it is also fast developing area with most of the country’s businesses located there. It was therefore not surprising that more than fifty percent of the reported cases were located in the business corridor along the urban and peri-urban areas in the country. The major cities were found to be an attraction because of their socio-economic amenities such as shopping centres, restaurants, government and public offices among others. Traffic is therefore more dense in the high risk areas and it has been observed that human mobility is one of the major determinants of COVID-19 (Gilbert et al., 2020; Kraemer et al., 2020; Rader et al., 2020; Zheng et al., 2020). Therefore, efforts aimed at minimizing transmission from human contact (such as increasing hand sanitizing stations, cancelling large group events, and
dispersing information about best practices for reducing transmission) would be most important in the places associated with high transmission risk. In addition, the current worldwide adopted mitigation strategy of lockdowns which aimed to limit population movement should be more enforced in such places. There would also be the need for governments to prevent or discourage movement in and out of such places to avoid transporting infections into low risk areas.

The exposure risk map (Fig. 2) highlighted the importance of major populated areas as potential epicentres where exposure was found to be very high. The more the number of (contagious) individuals reported in that particular area, the greater the likelihood that more people will be infected in the absence of risk reduction strategies especially at individual level. These were also areas with potentially high numbers of super-spreaders such as daily commuters, public health personnel including fun seekers, particularly in congested cities and peri-urban areas where there would often be constant risk from crowding, business activity and other associated health hazards accumulation. Furthermore, the large number of relatively young asymptomatic individuals who had been found to be potentially contagious confounds the risk of spread in high density areas (Gloos et al., 2020; Yang et al., 2020). We found that most of the high exposure risk areas overlap with areas characterized by high transmission risk. This was partly expected as exposure is a prerequisite for transmission and where it is high transmission would also be higher than the rest of the country. Such areas were those that attract large numbers of people for employment and other economic opportunities such as trading (both formal and informal) among other activities associated with urban lifestyles. The country’s major shopping centres, lodges, restaurants and markets for goods were found in many of such areas. Interestingly, we noted that most of the country’s large and popular churches were also found in this area. Churches were singled out because of their role in public gatherings and crowds as congregants meet together for prayers. Churches were also associated with being among the early drivers of transmission in the initial detection of the disease. Therefore their location is key in determining potential for super-spreaders as it is common for people to travel long distances across towns and regions to attend church services every Sunday in the country.

The susceptibility risk map in Fig. 3 showed areas where morbidity and mortality rates could potentially be highest in Eswatini. For instance, areas closer to many reported cases would have a higher risk of exposure than those farther away. This would imply that people in high exposure risk areas have a higher probability of being in contact with known cases or their subsequent contacts. However, it should be noted that unknown or undetected cases present the most danger and, if such exist, may significantly increase the exposure risk currently identified. The map indicated that these were areas with high concentrations of the elderly and people living with chronic illnesses. These were also areas with relatively high levels of poverty, predominantly in the rural areas which would often be characterized by poor service provision compared to the urban parts of the country. This map is important because it demonstrates to policy makers which areas would need prioritization in the fight against COVID-19. Therefore preparation efforts
could be scaled-up in such areas to minimize the impact of the disease on the available healthcare facilities and also to avoid unnecessarily overwhelming the healthcare staff should the disease strike. These areas were also found to be susceptible because of poverty and inadequate incomes to even afford the healthcare service and other basic necessities such as a healthy diet to boost the immune system and fight off the

Fig. 4. COVID-19 resource scarcity risk map for Eswatini.

Fig. 5. Multivariate cluster box plot chart for the COVID-19 zones in Eswatini.
disease from the body especially in the absence of either a vaccine or a
cure. Notwithstanding the use of non-therapeutic mitigation measures
such as the basic hand washing with soap and running water, most of
such places do not have this luxury as water scarcity is rife in the rural
parts of the country and most homesteads rely on rainwater harvesting
which itself is seasonal.

We also mapped healthcare resource scarcity which was found to be
high in most parts of the country, except in the major urban areas
(Fig. 4). The analysis only considered the major healthcare facilities
which usually have the capacity to respond to COVID-19 cases through
referral, admission and on-call medical doctors. Hence, even though
numerous small clinics can be found in various parts of the country,
these facilities mostly do not have the required infrastructure to respond
to COVID-19 and some of their services were inadequate and often non-
existent in most cases. The healthcare resource risk map showed the
level of preparedness to deal with increases in COVID-19 cases among
the most vulnerable populations which was found to be mostly in the
rural areas of the country. If COVID-19 were to spread rapidly in
Eswatini, it is highly likely that existing healthcare resources would be
quickly overwhelmed and its capacity diminished in like manner.

Hence, there is a need to develop strategies and put in place mechanisms
to bring people closer to the appropriate healthcare facilities (including
quarantine facilities) and avail the necessary equipment (such as test
kits, ventilators, and protective clothing and masks). Out of a total
population of 1 093 238, the number of people located in each of the
identified risk levels is summarized in Table two where percentages of at
risk people are shown.

### 4.2. Risk profiles and zones

Three (3) risk clusters or zones were obtained from the multivariate
clustering analysis. The box plot chart contains three lines that corre-
spond to the clusters on the map (Fig. 5). The nodes on the lines indicate
the risk level for each category. The largest values (at the top of the box
plot chart) represent the highest risk. Table 3, which corresponds to the
box plot chart, shows the different number of people and households
that were located in each of the identified risk zones which may poten-
tially be infected with COVID-19 in Eswatini. The zoning could also be
used to implement the current government strategy of lockdowns by
ensuring that there should be no movement of persons in between these
zones especially from areas of heightened risk to low risk areas. This way
government could minimize onward transmission to the rest of the low
risk areas in the country.

Fig. 6 shows the different risk profiles and zones across the country as
well as the areas which could potentially be harbouring different
infection levels among the population. Zone 1 (Red) was a cluster
comprised of areas that were potential epicentres as well as urban areas,
peri-urban areas and other highly populated areas. Such areas were found to be least affected by healthcare resource risk, but have high exposure, susceptibility and transmission risks. Hence, this zone should also be the focus of intense screening, testing and contact tracing efforts as the country attempts to detect all COVID-19 cases. In addition, this zone should also be given priority for minimizing human interaction and social distancing, that is, keeping children home from school, limiting visits to areas where there are elderly, people with comorbidities and hospitals, banning major events, and encouraging work from home. An extreme mitigation action should be to restrict movement in and out of such areas. This would be necessary especially because this zone supports the major health facilities where the risk of healthcare workers exposure is also very high.

The Zone 2 (Orange) cluster had the second highest risk for insufficient resources, exposure, susceptibility and transmission risks. The areas within this zone were those immediately on the periphery of Zone 1 areas not far from the urban areas but located within rural areas. In addition to practicing social distancing, areas within this zone should limit spread by putting plans in place isolation facilities in addition to conducting intensive healthcare training. Social distancing measures would also be very key in controlling spread into and within this zone. Zone 3 (Green) was a cluster of areas that had the highest risk for insufficient resources but currently had the least risk in terms of exposure and transmission risk. Sparsely populated and remote rural areas dominated this zone. These were areas where every effort needs to be put in ensuring that the virus does not reach them and hence should be strictly protected. Incursions of the virus spread to this area, an insurmountable problem would emerge thereby requiring huge amounts of resources for any response. Entry of the virus into these areas would imply a nation-wide emergency that could devastate the vulnerable population residing therein.

The results showed that about 29% (320 818) of the population was located in the high risk area for COVID-19 and containment of the risk in this zone would be very crucial if the country aimed to prevent and stop a nationwide outbreak. The high risk zone also coincided with high population movement as seen from the high traffic densities in very close proximity to the red zone. The zone is also densely populated as it falls within the two major cities of Eswatini which are Mbabane and Manzini. These two cities provide employment opportunities for most of the country’s working class and daily movement between them is not uncommon. The major cities also draw people from the rural areas that travel for various reasons and thus they would pose a great risk of transporting the virus to the low risk areas. Although this zone was found to be least affected by healthcare resource risk, it had high exposure, susceptibility and transmission risks meaning that the threat of overwhelming the available health facility services were still high. Unless strategic check points are put in place and population movement is controlled, there are high chances that COVID-19 would easily spread into the rural areas of the country and consequently affect highly susceptible individuals, the consequences of which could be dire.

The second zone (orange) had about 36% (394 288) of the country population and had highest risk for insufficient resources, exposure, susceptibility and transmission risks. These were areas that could be quickly overwhelmed as soon as the pandemic reached them since they already suffered from resource insufficiency. In addition to resource insufficiency they also had high exposure since they were located in close proximity to the high risk red zone which could be the risk factor that fuelled onward nationwide transmission. The third zone (green) was found to suffer mostly from resource scarcity and the severity of the pandemic could quickly be felt in such areas as they lacked the most basic necessities required to ward off the virus albeit currently incurable. Planning response measures could be aided by these identified COVID-19 risk zones which have the potential to support and guide the surveillance and preparedness efforts by the country.

5. Conclusions

We have identified the risk areas in the country and quantified the populations that were located in each of the risk levels. Successful surveillance and control of COVID-19 would depend on evidence based information and our study had demonstrated that it is possible to identify and map at risk areas even in the midst of data scarcity. We used various socio-economic and demographic variables which were crucial determining factors for the disease epidemiology and spread in the population. Our study had shown that major cities would be main sources of COVID-19 disaster in the country and therefore minimizing movement in and out of these places could be key in the control of the disease. Highly populated areas remain potential epicentres for the disease to spread and our work could guide government in implementing its mitigation strategies involving measures such as lockdowns, crowd and movement control efforts inter alia. The work has also identified highly susceptible areas which could potentially be easily overwhelmed should the disease reach them. This was important for government to know which areas need to be prioritised with preparedness and response measures in the country.

Estimating and quantifying the approximate number of people who may potentially be infected with COVID-19 remains impossible under resource scarcity and limited capacity especially in sub-Saharan Africa. Our study has mapped the areas with the highest risk of resource scarcity and government could use these results to identify areas that would need strengthening in terms of healthcare service provision. A majority of cases and deaths attributed to COVID-19 would likely remain unknown in the country, however our study could be used to estimate the potential burden of the disease. Our analysis showed that a total of 320 818 people were located in the high risk red zone and these were people who could potentially be infected with COVID-19 in the absence of mitigation measures. We have provided an estimation method that could support the pandemic risk forecasting, preparedness and response measures in the midst of data scarcity. Our work could be extended to include risk mapping of other infectious diseases such as common cold and seasonal flu in the country. The resultant map products could be used to guide on-the-ground surveillance and response efforts.

Declaration of interest

All authors declare no financial and no personal relationships with other people or organizations that could inappropriately influence (bias) their work.

Author contribution

SND and SDM drafted the manuscript, WMD and SPS analysed the data and produced the map products. All authors read and approved the final manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.apgeog.2020.102358.

References

Bedford, J., Enria, D., Giesecke, J., Heymann, D. L., Ihekweazu, C., Kohinger, G., … Wieler, L. H (2020). COVID-19: Towards controlling of a pandemic. The Lancet, 395, 1015-1018. https://doi.org/10.1016/S0140-6736(20)30673-5

Central Statistics Office. (2018a). The 2017 population and housing census - preliminary results. Mbabane.

Central Statistics Office. (2018b). Eswatini household income and expenditure survey 2016/17 - key findings report. Mbabane.

Dong, X. C., Liu, L. M., Bai, J. Y., Liu, Z. Q., Zhou, P. H., Gao, L., Li, X. Y., & Zhang, Y. (2020). Epidemiological characteristics of confirmed COVID-19 cases in Tianjin. Zhonghua Liu Xing Bing Xue Za Zhi Zhonghua Liuxingbingxue Zashi, 41, 638-642. https://doi.org/10.3760/cma.j.cn112338-20200221-00146
ESRI. (2020). How similarity search works [WWW document]. URL https://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/how-similarity-search-works.htm .

Facebook Connectivity Lab and Center for International Earth Information Network - CIESIN - Columbia University. (2016). High resolution settlement layer (HRSI). Source imagery for HRSI © 2016 DigitalGlobe Accessed DAY MONTH YEAR. [WWW Document]. URL https://ciesin.columbia.edu/data/hrsi/#acknowledgements.

Gilbert, M., Pullano, G., Pinotti, F., Valdano, E., Poletto, C., Boelle, P. Y., D’Ortenzio, E., Yezdanpanah, Y., Eholie, S. P., Altmann, M., Gutierrez, B., Kraemer, M. U. G., & Colizza, V. (2020). Preparedness and vulnerability of african countries against importations of COVID-19: A modelling study. The Lancet, 395, 871–877. https://doi.org/10.1016/S0140-6736(20)30411-6

Hou, J., Hong, J., Bi, D., Dong, B., Chen, Y., Ward, M. P., Tu, W., Jin, Z., Hu, J., Su, Q., Kazory, A., Ronco, C., & McCullough, P. A. (2020). SARS-CoV-2 (COVID-19) and AKI. The Lancet, 395, 843–846. https://doi.org/10.1016/S0140-6736(20)30414-0

Plessis, L., Faria, N. R., Li, R., Hanage, W. P., Brownstein, J. S., Layan, M., Vespignani, A., Tian, H., Dye, C., Pybus, O. G., & Scarpino, S. V. (2020). The effect of international travel and trade on the COVID-19 pandemic: A modeling study. Travel Medicine and Infectious Disease, 102358 (2020) 102358. https://doi.org/10.1016/j.tmaid.2020.101626

Zheng, R., Xu, Y., Wang, W., Ning, G., & Bi, Y. (2020). Spatial transmission of COVID-19 in China: A nationwide population-based piecewise mathematical modelling study. medRxiv. https://doi.org/10.1101/2020.03.27.20045757. 2020.03.27.20045757.

Huff, D. L. (2003). Parameter estimation in the Huff model. ArchSim. 3.

Kazory, A., Ronco, C., & McCullough, P. A. (2020). SARS-CoV-2 (COVID-19) and intravascular volume management strategies in the critically ill. Baylor University. Medical Center Proceedings, 1–6. https://doi.org/10.1085/089982820.2020.1754700

Kraemer, M. U. G., Yang, C. H., Gutierrez, B., Wu, C. H., Klein, B., Pigott, D. M., du Plessis, L., Faria, N. R., Li, R., Hanage, W. P., Brownstein, J. S., Layan, M., Vespignani, A., Tian, H., Dye, C., Pybus, O. G., & Scarpino, S. V. (2020). The impact of human mobility and control measures on the COVID-19 epidemic in China. Science. https://doi.org/10.1126/science.abb4218

Li, X., Xu, S., Yu, M., Wang, X., Tao, Y., Zhou, Y., Shi, J., Zhou, M., Bu, B., Yang, Z., Zhang, C., Yue, J., Zhang, Z., Renz, H., Liu, X., Xie, J., Xie, M., & Zhao, J. (2020). Risk factors for severity and mortality in adult COVID-19 inpatients in Wuhan. The Journal of Allergy and Clinical Immunology. https://doi.org/10.1016/j.jaci.2020.04.006

Lloyd, S. P. (1982). Least squares quantization in PCM. IEEE Transactions on Information Theory, 28, 129–137. https://doi.org/10.1109/TIT.1982.1056498

Manuwa, S., Tsoka-Gwevheni, J., & Sartorius, B. (2019). The economic burden of HIV and type 2 diabetes comorbidity: Implications for care in countries with high burden of HIV. Journal of Diabetology, 10(3), 89–96. https://doi.org/10.4103/jod.jod_50_18

Ministry of Health. (2020). Coronavirus disease 2019 (COVID-19) and travel medical advice for travelers in Swaziland. https://doi.org/10.3201/eid2108.150665

Wilder-Smith, A., Chiew, C. J., & Lee, V. J. (2020). Can we contain the COVID-19 outbreak with the same measures as for SARS? The Lancet Infectious Diseases, 20, 1016–1017. https://doi.org/10.1016/S1473-3099(20)30129-8

World Health Organization. (2020). Modes of transmission of virus causing COVID-19: Implications for IPC precaution recommendations. Scientific brief, 1–3. https://doi.org/10.1056/NEJMoa2001215

Wu, Y.-C., Chen, C.-S., & Chan, Y.-Y. (2020). The outbreak of COVID-19: An overview. Journal of the Chinese Medical Association, 83, 217–220. https://doi.org/10.1097/JCMA.0000000000000270

Yang, K., Wang, L., Li, F., Chen, D., Li, X., Qiu, C., & Chen, R. (2020). Analysis of epidemiological characteristics of coronavirus 2019 infection and preventive measures in shenzhen China: A heavy population city. https://doi.org/10.1101/2020.02.28.20028555. medRxiv