

USING SOFTWARE TO EXTRACT KEY FINDINGS FROM SCIENTIFIC RESEARCH PAPERS: A CASE STUDY USING RESEARCH ON COVID-19

Jai Agarwal, Taruna Agarwal, Hrithik Jain, Yilun Wu, John Leddo, Saharsh Ranga, Riya Srikumar, Antonia Gabriel, Ritika Bishen, Emma Wu, Srikaran Yelimati, Max Ritter, Sophia Nasibdar, Eshanth Penumatsa, Ishwarya Ramineni, Eshan Iyer, Krishna Dotiwalla, Srihitha Somavarapu, Nishaaj Khan, Sia Magoon, Karthik Pathakota

Dr. John Leddo is the director of research at MyEdMaster.

MyEdMaster, LLC., Herndon, Virginia, USA

DOI: 10.46609/IJSSER.2022.v07i09.029 URL: <https://doi.org/10.46609/IJSSER.2022.v07i09.029>

Received: 27 September 2022 / Accepted: 5 October 2022 / Published: 10 October 2022

ABSTRACT

With the explosion of information on the Internet, search engine users still find themselves having to weed through a myriad of websites to ensure that they find the relevant information. This is even more cumbersome in dynamic subject areas, such as scientific research, where research findings may not be stable and even contradictory. Laypeople are especially burdened since they may lack the knowledge to evaluate what scientific papers are actually concluding. The present paper describes software that reads scientific papers and distills their principal findings in a format that laypeople can understand. This software is evaluated in the topic area of research on Covid-19. A separate paper evaluates this software in the topic area of research on diet.

Introduction

Over time, there has been an explosion of information available to people over the Internet. While search engines have become more sophisticated at retrieving information, virtually any search term a person enters into a search engine yields millions of websites. This creates inherent problems. No search engine user will have the time and inclination to go through the retrieved websites to determine which websites contain the best and most relevant information. Even within a given website, it is not always immediately apparent where the relevant information is. Of course, once the information is found, it may be difficult for the user to understand, particularly if the topic in question is technical in nature. Perhaps no subject area exemplifies this point more than medicine.

In our previous work (Boina et al, 2021a, b, c, d), we argued that search engines and personal assistants of the future need to do more than retrieve websites but also read through them and extract and learn information necessary to answer users' questions. In these studies, we developed software that accepted medical conditions from users, went on the Internet, retrieved information from the sites, learned it and then diagnosed medical conditions in the users with greater than 90% accuracy. These medical conditions included, among others, a variety of skin ailments and stress.

One of the characteristics that the medical conditions described in our previous work had in common is that the knowledge about these conditions was fairly stable, and that there were agreed upon criteria for diagnosing them. However, this is not true for all medical topics. For example, when Covid-19 came out, medical professionals scrambled to get a handle on this new disease, how it was transmitted, how lethal it was, how to cure it.

As with any new field of inquiry, such investigation still involves tracking a moving target. When Covid-19 first broke out, medical professionals viewed it as transmitted by surface contact. People were advised to wear gloves, they avoided shaking hands, and even were told not to touch their faces. Subsequently, Covid-19 was viewed as an airborne disease and mask mandates were implemented.

It is challenging enough for professionals to keep up with an evolving research base, but what if you are a layperson who just wants to keep healthy or cope with a Covid-19 infection? To address a problem such as this, it would be beneficial to have software that can read scientific journals and then summarize the key findings such that laypeople (and even professionals) have a tool that helps them weed through the complexities of virtually unlimited data and weed out the relevant findings. The purpose of the present project is to create such a tool.

The software

The code performs automatic information extraction from academic journal articles to the literature review template with certain specifications. This information extraction script is written in Python and utilizes natural language processing and text analysis. Applying this script saves significant time compared to manually reading journal articles and filling out information based on these articles. Note that this script only fills out partial columns of the template since the other information required to extract from the journal articles and papers earns high variance in context and needs subjective judgment from readers. The columns that will be filled out using this script: Title, Paper source (website), Year of Publication, Journal Type (review), Link, Human trial or not, Data (Interventions), Positive/negative findings, Conclusion Summary.

We use the script by keeping the code, the academic journal article pdf files, and the literature review template into the same directory. We then record all the file names and store them into a Python list under the “if __name__ == ‘__main__’:”. The first thing that happens after this is the extraction of DOI (Digital Object Identifier), which is a string of numbers, letters and symbols used to permanently identify an article or document and link to it on the web.

It is extracted by finding the start and the end index of the string which represents the doi and returns the string. This string can sometime contain a period in the end. That period is removed by removing the last character. Also the doi might start with an “org/”, so we ignore the first 4 characters. The last character in the doi should be a number, so we remove the last non-numeric characters unless a number is encountered and then return the string. This returned string is the DOI. Now, the DOI is converted to json using the doi2json function so that the query function can form the URL and then use the GET method to retrieve information from the given server using that URL. From this retrieved information, we can extract the title, publisher, year and journal type. After this, we find out the key information in the journal article. The approach taken here is to analyze the sentences between the “keywords” paragraph and “introduction” paragraph in the paper. Each word in the sentence between these paragraphs is matched with the values of the dictionary interventionWordBank. Wherever a match is found, its corresponding key of the value is returned. If no match is found, then ‘others’ is returned. The next thing to do is to find the effects of the experimental study (beneficial, detrimental or neutral). Sentiment analyses is performed on the paragraph that lies between the Abstract and the Keywords, if the label of the analyses is positive, then beneficial is returned. If negative then detrimental is returned, else neutral is returned. This is done by getting the start index of the Abstract and the end index of the Abstract. If the difference between the start and the end index is more than 512 words, then only 510 words starting from the start index is taken. Finally, we summarize the paper by analyzing the text between the heading Conclusion and Acknowledgements (if Acknowledgements are not there in the paper, then we analyze until References). This is done by finding the differences of first instance of the end index and the first instance of the words “Conclusion”, “Conclusion ”, “Conclusions ”, “Summary”, “Discussion”, “CONCLUSIONS”, “CONCLUSION”. The smallest of all these differences is taken as input to generate the summary. A dictionary of word-frequency pair is created where words that are not part of STOP_WORDS and punctuations are included with their value being the frequency of those words. After this, a dictionary of sentence-score pair is created, the score of the sentences is calculated by checking the number of words that the sentence and the word-frequency dictionary have in common, and then the frequency of that common word in the word-frequency pair dictionary is taken/added as score in the sentence-score pair dictionary. Then, the summary is calculated by finding the nlargest scores of the sentence where n here is the number of sentence

times the percentage of paragraph we want as summary. After getting all these values, the columns that can be filled in the template will be filled.

Testing the Software

The goal of the present software is to read scientific articles and write summaries of them that are understandable by laypeople, thus enabling them to keep up with the latest in scientific research. Accordingly, we tested the accuracy of the software by having it read scientific articles and write summaries. The summaries focused on the main relationship between the independent variables and the dependent variables. These summaries were compared with actual text taken from the scientific articles themselves. A total of 11 articles, each focusing on research involving Covid-19, were used. In a separate paper, we evaluate the software's effectiveness in reading and summarizing papers involving dietary interventions.

Results

The descriptions of text from the reviewed articles and the summaries produced by the software are shown below.

Title:- Simultaneous ventilation in the Covid-19 pandemic. A bench study

Link:- <https://doi.org/10.1371/journal.pone.0245578>

Conclusion from paper:-

The lower-level ventilator performed closely to the ICU ventilator. Due to dependence of VT to C pressure control should be used to maintain adequate VT at least in one test-lung when compliance and/or R change abruptly and monitoring of VT should be done carefully.

Conclusion from model:-

The lower-level ventilator performed closely to the ICU ventilator. Due to dependence of VT to C pressure control should be used to maintain adequate VT at least in one test-lung when compliance and/or R change abruptly and monitoring of VT should be done carefully.

Title:- Indoor Air Quality Strategies for Air-Conditioning and Ventilation Systems with the Spread of the Global Coronavirus (COVID-19) Epidemic: Improvements and Recommendations

Link:- <https://doi.org/10.1016/j.envres.2021.111314>

Conclusion from paper:-

The spread of health problems inside various buildings in the 1980s led to a review of prevention strategies from the World Health Organization and its emphasis on the importance of indoor air quality and its consideration of the elements that harm human health on the long term. This is because of the rise in temperatures due to global warming, the increase in air pollution rates and its impact on health, and the wide-spread use of air-conditioners of all kinds, especially if they are used incorrectly. Accordingly, the need for thoughtful foundations and recommendations to mitigate the severity of the global COVID-19 epidemic emerged as a result of total reliance on air-conditioning and heating systems, which may contribute to the transmission/spread of airborne diseases, as has been proven in the past in describing Japan, Germany and the Diamond Princess Cruise for SARS. Moreover, the epidemiological characteristics of SARS-CoV-2 are not yet clear, and most recent research has indicated that SARS-CoV-2 spreads faster in winter than in summer, which is an indication of the importance of temperature and humidity in the transmission of COVID-19. The current research lists previous reviews of multiple air pollutants and the various approved and modern methods of purification and filtering methods with an indication of the impact of temperatures and humidity on the spread of the global epidemic of COVID-19; besides, a review of the different air conditioning systems is included to adopt considerations and recommendations to limit and prevent the spread of the global COVID-19 virus as much as possible according to the following guideline recommendations: 1. To meet the needs of indoor air quality, it is necessary to take into account that UVGI is only a primary filtration stage; although it is germicidal, but it is not capable of killing or inactivating infectious microorganisms such as coronavirus. 2. Avoid recirculating air in HVAC systems by closing the return air dampers by diverting the return air of central air conditioning systems to the exhaust air path and feeding 100% of total fresh air. 3. The exhaust air path must be at a height of at least 5 m from the end of the top of the building and the size of the exhaust fans must increase. 4. Installing independent, decentralized air-conditioning units in the isolation rooms for COVID-19 sufferers and closing the central air-conditioning systems in these rooms. 5. HEPA air filters in air conditioners should be replaced by nanofibrous air filters or enhanced electrostatic air filters. 6. Air conditioners that do not have fresh air require opening the windows partially to renew the air in places to provide the largest possible amount of outside air. 7. Negative pressure must be maintained in ventilation systems for isolation rooms for COVID-19 infected patients, as well as in the accompanying bathrooms, while the corridors must be kept at positive pressure. 8. Reducing levels of air pollution inside polluted places and cities, especially industrial ones, by implementing strict, compulsory, and sustainable environmental policies at the international level. 9. Close the path between the corridors and the residents' rooms in central air-conditioning systems in the applications of hospitals and hotels etc., that are designed to draw air from multi-zones air handling units through the suspended ceiling to maintain the negative pressure difference. 10. The amount of exhaust air should be adjusted greater than the supply air through

the central A/C systems so that a negative pressure of at least 2.5 Pa (preferably > 5 Pa) is achieved in the zones. 11. The amount of air inside must be changed at least 12 times per hour, with a carbon dioxide level at 400 parts per million to increase the delivery of the fresh air. 12. To quell the COVID-19 virus spreading rate, set the conditioned room temperature between 25 °C and 27 °C while maintaining the relative humidity between 50% and 70%. 13. For treating the air in special cases, the suspended air with bacteria and viruses must be exposed to ultraviolet radiation (15–20 min) and raising its temperature (45 min at 75 °C). 14. Maintaining higher ventilation rates for a longer period with the people staying in the places, whilst set the ventilation rates at a lower level when people are absent at low velocities to reduce the consumed power. 15. Apply MERV 7 as primary filtration and MERV 14 as secondary filters to remove 98% of airborne particles with a diameter of 0.3–1.0 mm. 16. Periodically disinfect HVAC systems, including air ducts, outlets, fans, and air conditioning and heating equipment using chemical disinfection (1% hypochlorite). 17. Avoid the use of circulation fans (flipping fans) in public places inhabited by people. If necessary, the air should be renewed by increasing the flow of outside air (fresh air). 18. Avoid setting air conditioning control systems to low temperatures (below 70 F/21 C) and low humidity settings (less than 40%) to reduce the survival time of SARS-COV-2 in an indoor environment. 19. Commitment to regular maintenance procedures of HVAC equipment and the necessity to wear appropriate personal protective equipment and wrap all damaged materials, such as old filters, and dispose them safely by burning (OSHA 3990–03 2020). 20. Using a strategy of a proper ventilation system, effective air purification technology, humidity regulation, and temperature control that improve indoor air quality and protect against airborne infectious diseases, especially COVID-19.

Conclusion from model:-

“Indoor air quality strategies for air-conditioning and ventilation systems with the spread of the global Coronavirus (COVID-19) epidemic: Improvements and recommendations.” made with individual efforts to be important, novel in its presentation of the topic and not published before, and it benefits all humanity as a whole, due to its connection to the topic of the day worldwide (COVID-19). Avoid recirculating air in HVAC systems by closing the return air

dampers by diverting the return air of central air conditioning systems to the exhaust air path and feeding 100% of total fresh air. Humidity on the spread of the global epidemic of COVID-19; besides, a review of the different air conditioning systems is included to adopt considerations and recommendations to limit and prevent the spread of the global COVID-19 virus as much as possible according to the following guideline recommendations:

1. Close the path between the corridors and the residents' rooms in central air-conditioning systems in the applications of hospitals and hotels etc., that are designed to draw air from multi-

zones air handling units through the suspended ceiling to maintain the negative pressure difference. Accordingly, the need for thoughtful foundations and recommendations to mitigate the severity of the global COVID-19 epidemic emerged as a result of total reliance on air-conditioning and heating systems, which may contribute to the transmission/spread of airborne diseases, as has been proven in the past in describing Japan, Germany and the Diamond Princess Cruise for SARS. Conclusions and guideline recommendations

The spread of health problems inside various buildings in the 1980s led to a review of prevention strategies from the World Health Organization and its emphasis on the importance of indoor air quality and its consideration of the elements that harm human health on the long term.

Using a strategy of a proper ventilation system, effective air purification technology, humidity regulation, and temperature control that improve indoor air quality and protect against airborne infectious diseases, especially COVID-19.

“Indoor Air Quality Strategies for Air-Conditioning and Ventilation Systems with the Spread of the Global Coronavirus (COVID-19) Epidemic: Improvements and Recommendations”. The amount of exhaust air should be adjusted greater than the supply air through the central A/C systems so that a negative pressure of at least 2.5 Pa (preferably > 5 Pa) is achieved in the zones.

This is because of the rise in temperatures due to global warming, the increase in air pollution rates and its impact on health, and the widespread use of air-conditioners of all kinds, especially if they are used incorrectly. Air conditioners that do not have fresh air require opening the windows partially to renew the air in places to provide the largest possible amount of outside air. Commitment to regular maintenance procedures of HVAC equipment and the necessity to wear appropriate personal protective equipment and wrap all damaged materials, such as old filters, and dispose them safely by burning (OSHA 3990–03 2020). Moreover, the epidemiological characteristics of SARS-CoV-2 are not yet clear, and most recent research has indicated that SARS-CoV-2 spreads faster in winter than in summer, which is an indication of the importance of temperature and humidity in the transmission of COVID-19. Avoid setting air conditioning control systems to low temperatures (below 70 F/21 C) and low humidity settings (less than 40%) to reduce the survival time of SARS-COV-2 in an indoor environment.

Title:- A review of methods to reduce the probability of the airborne spread of COVID-19 in ventilation systems and enclosed spaces

Link:- <https://doi.org/10.1016/j.envres.2021.111765>

Conclusion from paper:-

Several recommendations may be made based upon current research and published guidelines. In summary, a layered approach that combines several of the proposed methods is recommended. This is because none of the individual methods are likely to completely remove the viral airborne particles from an enclosed space, and a layered approach may potentially avoid the pitfall of diminishing returns from a single method of viral removal. The ideal approach may be unique for each situation to ensure the maximum removal of airborne particles containing COVID-19 and would depend on factors such as the ability to introduce fresh air to a space, or the configuration of a previously installed HVAC system. It is worth noting again that the conceptual basis for the different technologies discussed above to increase the viral particle removal rate is predicated upon several simplifying assumptions. Namely, these are assumptions surrounding the infectivity and survivability of COVID-19, and that an enclosed space is well-mixed. The infectivity and survivability of COVID-19 are directly related to the concept of a quantum, and the amount of the virus that a healthy individual can ingest before becoming infected as well as the time that the airborne virus may be G. Berry et al. *Environmental Research* 203 (2022) 111765 9 exposed before inactivation. The well-mixed room assumption simplifies predictive models by assuming that an enclosed space has a uniform distribution of viral particles, but in doing so ignores the realistic dynamic effects induced by proximity to an infected individual. Table 1 suggests advantages, disadvantages, and research gaps for each category of methods. In summary of the individual methods:

- Ventilating an enclosed space provides an effective way to reduce the concentration of airborne particles carrying the COVID-19 virus. This performance of an already-in-place HVAC system less than a HEPA filter will, or some other filter that will induce a larger pressure drop.
- Air ionization provides a promising method to reduce the number of airborne particles by increasing their deposition rate to walls and other surfaces. However, this does not necessarily inactivate the virus, and the argument can be made that this may increase the chance of transmitting the virus through contact. A potential disadvantage related to air ionization is the generation of ozone. However, the mechanisms of ozone generation are understood relatively well for this method, and there exist options to mitigate an unsafe amount from being introduced into an enclosed space.
- Environmental controls show promise to create an environment that is adverse to the survival of the virus. Cool and dry environments are known to increase the chances of infection, where it is suggested that a cool environment increases the survival rate of the virus, and a dry environment causes the airborne droplets to evaporate to a higher degree, thus causing them to remain airborne for a longer period of time and increasing the survivability of the virus. To combat this, a warmer environment is recommended, but equally important is a range of humidities that will maintain comfort, but also reduce unnecessary and excessive evaporation of the airborne droplets.
- UVGI offers a promising solution to inactivate viruses, which has been proven to work in other situations similar to the current situation with COVID-19. It may be installed in existing HVAC systems, or the upper

portion of a room to reduce the exposure of people to possibly harmful radiation at sufficient doses. Some potential disadvantages of UVGI technology are the possible generation of ozone and the degradation of materials exposed to ultraviolet radiation. However, using glass to absorb the wavelength of light responsible for ozone generation presents a potentially effective way to address this issue. Also, using reflective materials in the duct not only circumvents material degradation, but it will also improve the efficacy of viral inactivation. • Nanoparticles offer a solution to increase the efficiency of a fibrous filter while also providing a method to inactivate captured viral particles. However, the addition of nanoparticles to filter media also increases the pressure drop across the filter, increasing the impact of implementing this technology to an existing HVAC system. • Non-thermal plasma has been used to inactivate bacteria and various airborne pathogens including the MS2 bacteriophage with varying levels of success. Like UVGI, it is possible to install a portable air sterilizing device in a room or a duct to inactivate a virus. However, since NTP relies on the generation of ROS, unacceptable levels of ozone may likely be generated. This may be mitigated depending on the form of NTP generation, such as an MPCD reactor, or by using an ozone filter. • Chemical disinfectants have been suggested and used to inactivate viruses in enclosed spaces, and research exists suggesting that their implementation is highly effective. Some of the disinfectants discussed are not harmful to humans, but more research is required to quantify the effectiveness of these methods in addition to the safety concerns regarding an application with HVAC systems. • Using an elevated air temperature to inactivate the virus has also been proven to be effective. It is possible to expose the viral particles to a small-scale heat source, such as an axial heating element or a heated metal filter in a duct, offering a possible solution to the inactivation of viral particles in an HVAC system. Furthermore, using combustion to generate halogens in an airstream is effective at deactivating viral particles while also reducing the amount of heat added to a system. However, the introduction of additional pollutants from combustion, and additional heat to a system from the heat inactivation method in general, is undesirable. • There is an observed lack of useable cost analysis data regarding the implementation of the technologies presented. • The reported efficiencies and removal rates of the presented technologies are not necessarily directly comparable to each other. For example, the removal rate of a mechanical filter is dependent on the filter's efficiency and the supplied room's volume and airflow rate, while several of the reported values for the efficiency of air ionization are from experiments conducted in sealed enclosures with no air circulation

Conclusion from model:-

- Proven and reliable inactivation method with technologies that can be readily integrated into HVAC systems.
- Relatively easy to incorporate coated filters into an existing HVAC system.

- Simplifying assumptions regarding the dynamics of the viral transmission will affect the perceived effectiveness of simple methods, such as opening a window or promoting airflow by utilizing fans. Cool and dry environments are known to increase the chances of infection, where it is suggested that a cool environment increases the survival rate of the virus, and a dry environment causes the airborne droplets to evaporate to a higher degree, thus causing them to remain airborne for a longer period of time and increasing the survivability of the virus.
- The nanoparticle field is currently growing and will continue to expand in applications and available technology.
- Long-term use of high-powered ionizers can potentially cause a buildup of charge in a small, enclosed space.

For example, the removal rate of a mechanical filter is dependent on the filter's efficiency and the supplied room's volume and airflow rate, while several of the reported values for the efficiency of air ionization are from experiments conducted in sealed enclosures with no circulation.

- Commercially available configurations that may be applied to individual enclosed spaces, or entire HVAC systems. Ultra-Violet Germicidal Irradiation (UVGI) does not actively remove or inactivate viral particles as compared to the other technologies.

It is possible to expose the viral particles to a small-scale heat source, such as an axial heating element or a heated metal filter in a duct, offering a possible solution to the inactivation of viral particles in an HVAC system.

- The Nickel-foam filter media and heat inactivation mask require more research to determine efficacy and reliability.
- The currently available coatings are not as efficient as other methods of viral inactivation. However, the mechanisms of ozone generation are understood relatively well for this method, and there exist options to mitigate an unsafe amount from being introduced into an enclosed space.
- The effects of filter loading have not been thoroughly investigated regarding the longevity of inactivation efficiency.

Title:- Natural ventilation strategy and related issues to prevent coronavirus disease 2019 (COVID-19) airborne transmission in a school building

Link:- <https://doi.org/10.1016/j.scitotenv.2021.147764>

Conclusion from paper:-

Field measurements were used to analyze natural ventilation strategies using window opening in a school building to cope with the current COVID-19 situation. The infection rate was calculated according to the natural ventilation performance, exposure time, and whether a mask was worn. This study suggests an appropriate window opening area to prevent viral air infection that can be applied in real-world conditions. The main results of this study are as follows: (1) The ventilation rates according to window opening ratios of 15, 30, and 100% under cross-ventilation were measured at 6.51, 11.20, and 22.43 ACH, respectively, in the summer season in Korea. (2) The infection possibility can be maintained at less than 1% by securing 6.51 ACH (1093.7 CMH), restricting exposure time to less than 3 h and wearing a mask. The Korean government's guidelines for opening windows 30% were found to provide a sufficient natural ventilation rate to prevent airborne virus infection in public buildings. (3) Wearing a mask can reduce the infection probability by four times compared to not wearing it, and the infection probability increased arithmetically as the exposure time increased. In situations where sufficient ventilation is not available, it is important to wear a mask or minimize exposure time. (4) Power consumption by an air conditioner increased by 10.2% and 22.5% under opening ratios of 15% and 30%, respectively compared to conditions where all windows were closed and there was no natural ventilation. Even when air conditioning and natural ventilation occur at the same time, the additional cost is not high in the summer in Korea. (5) Cross-ventilation is efficient in terms of air exchange rate compared to single-sided ventilation, and cross-ventilation is recommended to minimize the infection possibility in high-density public buildings. If cross-ventilation is not possible, it is advisable to use an auxiliary fan to achieve the same effect as cross ventilation.

Conclusion from model:-

Even when air conditioning and natural ventilation occur at the same time, the additional cost is nothigh in the summer in Korea.The main results of this study are as follows:

Cross-ventilation is efficient in terms of air exchange rate compared to single-sided ventilation, and cross-ventilation is recommended to minimize the infection possibility in high-densitypublic buildings.

The ventilation rates according to window opening ratios of 15,30, and 100% under cross-ventilation were measured at 6.51,11.20, and 22.43 ACH, respectively, in the summer season inKorea.

The infection possibility can be maintained at less than 1% by securing 6.51 ACH (1093.7 CMH), restricting exposure time to lessthan 3 h and wearing a mask.

Title:- Learning from past respiratory failure patients to triage COVID-19 patient ventilator needs: A multi-institutional study

Link:- <https://doi.org/10.1016/j.jbi.2021.103802>

Conclusion from paper:-

Medical knowledge often progresses by analogizing from pathophysiologically related but distinct disease entities. To address the lack of COVID specific data, we used the abundant historical data from related conditions to address prognostication for a novel epidemic infection. The commonality between COVID-19 and other common respiratory failure diseases can help provide guidance to providers treating COVID-19 patients during this novel pandemic where clinical experience and evidence remain sparse. When validated on COVID-19 patients from two healthcare systems, our best-performing model demonstrated the capacity to meaningfully inform the clinical suspicion that a COVID-19 patient will need IMV. Future efforts to increase model sensitivity and further bolster positive predictive value will improve the model's clinical utility.

5.1. Lessons learned in the beginning of the pandemic, many institutes were challenged with a limited number of cases for model training and limited knowledge about the novel COVID-19 disease. The quick surge of COVID-19 patients overflows health facilities and a decision support tool was in great need for triaging patients. Many rushed to produce models with limited training data [20] that led to potentially biased models [24]. This work proposes an alternative approach that addresses the needs of many institutions to support decision making in the void of robust data sets. Such a framework could support the rapid dissemination of prognostic models at the next pandemic or for rare diseases. In the haste of developing AI-based COVID decision tools, most published models have fallen short of providing safe and effective guidance [24]. In a continuously updated systematic review of COVID-19 prediction models, most models lack transparency and are at high risk of bias [20]. External validation is not common and performance drops dramatically when these models are tested on a different dataset, which can make these models irrelevant for deployment at the point of care outside the training data site [26]. The model reported in this study also suffers from generalizability to some extent, however there are important lessons that can be learned from this endeavor. Data extraction procedures vary greatly across systems. To ensure generalizability, one must empirically validate each feature used in the model, including missingness, distribution, time captured, and units captured. For example, oxygen flow was an essential feature of this model, yet captured very differently at the two sites. For one site, conversion of values for FiO₂ were performed; the values recorded included either the liters of oxygen per minutes delivered or the actual FiO₂ values, we converted the oxygen delivered into FiO₂ when necessary. While the other site has this feature recorded cleaner and more complete. In this work, one institute created a panel to obtain many features for each COVID patient hospitalized. In contrast, features at the

other hospital had to be scraped from flowcharts. Generalizability was also affected by the differences in patient populations, where there were important differences in race, ethnicity, smoking status, and pre-existing conditions, which might also be a reflection of data capture. However, given that we are one of the few COVID models that have actually performed cross-site validation, the model is likely more generalizable than most. This work demonstrated the feasibility of cross-site validation in a very short period of time and suggests that clinical decision tools need to prove robustness against a wide range of generalizability challenges. In addition to model development, we provide guidance on the implementation of our framework and multi-institute validation. First, it is important to identify ‘like-cohorts’ that share outcome characteristics and clinical manifestations with the disease of interest. These like cohorts must be common with sufficient sample sizes for training, testing, and validating the ML models. Second, a system should be in place to extract cohorts and variables from the EHRs and, importantly, identify new variables that might be unique to the disease of interest and stored in flowsheets or as unstructured data in clinical narrative text. Third, at the onset of the pandemic, COVID-19 treatment was unsure and guidelines did not exist, large practice variation was seen across the globe and needs to be considered when implementing clinical decision support tools. Therefore, cross-validation across care settings is necessary. During cross-validation, every variable used in the model must be thoroughly investigated, including data capture, storage, missingness and validity. Finally, geographical variation must be considered, as the outbreak of COVID-19 in New York City is greatly different from outbreaks in other cities across the globe, again highlighting the need for cross validation. The lessons learned for clinical model application were that models trained on data of a “patient-like me” cohort (i.e., COVID-like diseases) can have good accuracy and present clinical utility by identifying in patients at high risk of decomposition. Importantly, we also learned that laboratory factors associated with inflammation (i.e., D-dimer, ferritin, LDH, CRP), that have not been previously appreciated as predictors for patient deterioration in respiratory illnesses, were highly influential in our model. The framework provides an opportunity to quickly identify H. Carmichael et al. *Journal of Biomedical Informatics* 119 (2021) 103802 8 unsuspected risk factors associated with disease outcomes at the onset of future pandemics.

Conclusion from model:-

A system should be in place to extract cohorts and variables from the EHRs and, importantly, identify new variables that might be unique to the disease of interest and stored in flowsheets or as unstructured data in clinical narrative text. Importantly, we also learned that laboratory factors associated with inflammation (i.e., D-dimer, ferritin, LDH, CRP), that have not been previously appreciated as predictors for patient deterioration in respiratory illnesses, were highly influential

in our model. The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

The lessons learned for clinical model application were that models trained on data of a “patient-like me” cohort (i.e., COVID-like diseases) can have good accuracy and present clinical utility by identifying inpatients at high risk of decomposition. External validation is not common and performance drops dramatically when these models are tested on a different dataset, which can make these models irrelevant for deployment at the point of care outside the training data site [26]. This work demonstrated the feasibility of cross-site validation in a very short period of time and suggests that clinical decision tools need to prove robustness against a wide range of generalizability challenges. Finally, geographical variation must be considered, as the outbreak of COVID-19 in New York City is greatly different from outbreaks in other cities across the globe, again highlighting the need for cross validation.

To address the lack of COVID specific data, we used the abundant historical data from related conditions to address prognostication for a novel epidemic infection. When validated on COVID-19 patients from two healthcare systems, our best-performing model demonstrated the capacity to meaningfully inform the clinical suspicion that a COVID-19 patient will need IMV. Third, at the onset of the pandemic, COVID-19 treatment was unsure and guidelines did not exist, large practice variation was seen across the globe and needs to be considered when implementing clinical decision support tools. The commonality between COVID-19 and other common respiratory failure diseases can help provide guidance to providers treating COVID-19 patients during this novel pandemic where clinical experience and evidence remain sparse.

Title:- Particle modeling of the spreading of coronavirus disease (COVID-19)

Link:- <https://doi.org/10.1063/5.0020565>

Conclusion from paper:-

This paper presented a kinetic Monte Carlo algorithm for modelling different scenarios of the infection rate of the novel coronavirus disease. This model’s main feature lies in its extreme flexibility and in the fact that the parameter R_0 is obtained from the simulation and not pre-assumed. It can rather be used, in principle, as a way to tune up the other parameters better based on the post-processing of clinical and epidemiological data. Although it is challenging to model the specific characters of each coronavirus infected area, our results show that strict social distancing and a cyclic time pattern might help to keep the infection rate under control over a long period, even for an intrinsic doubling time of 2.5 days and in the presence of infection from unknown sources. Our ability to model and prove differences between the different lockdown patterns sharpens the need for physical and mathematical models that allow examining different

ways for reducing the spread of the epidemic. From the physical point of view, effective strategies for controlling the infection rate of a specific area should lower its effective temperature as much as possible by keeping social distancing and avoiding creating hot spots such as those related to high concentrations of people on a daily basis.

Conclusion from model:-

Although it is challenging to model the specific characters of each coronavirus infected area, our results show that strict social distancing and a cyclic time pattern might help to keep the infection rate under control over a long period, even for an intrinsic doubling time of 2.5 days and in the presence of infection from unknown sources.

Title:- Effectiveness of non-pharmaceutical public health interventions against COVID-19: A systematic review and meta-analysis

Link:- <https://doi.org/10.1371/journal.pone.0260371>

Conclusion from paper:-

The majority of NPHIs had positive effects on restraining the COVID-19 spread. We found significant decreases in COVID-19 case growth rate, death growth rate, and reproduction number during and in the later stage of the lockdown. However, it was challenging for countries to maintain this path after the lockdown was lifted. The early enforcement of lockdown, when the incidence rate is not high, can contribute to a shorter duration of lockdown and a lower increase of the case growth rate in the post-lockdown era. Considering the negative impact of the nationwide lockdown and other harsh restrictions on the economy and people's life, such interventions should be mitigated by adopting other NPHIs such as mass mask-wearing, patient/suspected case isolation strategies, and contact tracing. With the problems that remain regarding universal access to vaccines and their effectiveness, more public health strategies are needed not only to flatten the epidemic curve but to maintain it flat as well. This is particularly important when thinking of preparing for upcoming pandemic waves and future epidemics. The results of this systematic review and meta-analysis could aid policymakers in planning future public health interventions and is vital to understand the impact of NPHIs adopted by countries. PLOS ONE Systematic review of effectiveness of non-pharmaceutical public health interventions against COVID-19 PLOS ONE | <https://doi.org/10.1371/journal.pone.0260371> November 23, 2021 13 / 19 Although almost all studies showed significant, positive effects of NPHIs on controlling the COVID-19 pandemic, we cannot dismiss the side impact of stringent restricting interventions on people, such as compliance with medications, healthy lifestyle habits among patients with non-communicable disease, delay in cancer diagnosis, and mental health problems. Further research using a wide range of covariates are needed to evaluate the

effectiveness of NPHIs in different countries. Studies need to address the impact of NPHIs on the population's other health problems than COVID-19 as well.

Conclusion from model:-

Although almost all studies showed significant, positive effects of NPHIs on controlling the COVID-19 pandemic, we cannot dismiss the side impact of stringent restricting interventions on people, such as compliance with medications, healthy lifestyle habits among patients with non-communicable disease, delay in cancer diagnosis, and mental health problems. Systematic review of effectiveness of non-pharmaceutical public health interventions against COVID-19

The early enforcement of lockdown, when the incidence rate is not high, can contribute to a shorter duration of lockdown and a lower increase of the case growth rate in the post-lockdown era. Considering the negative impact of the nationwide lockdown and other harsh restrictions on the economy and people's life, such interventions should be mitigated by adopting other NPHIs such as mass mask-wearing, patient/suspected case isolation strategies, and contact tracing. The results of this systematic review and meta-analysis could aid policymakers in planning future public health interventions and is vital to understand the impact of NPHIs adopted by countries.

Results of the EPOK risk of bias assessment for studies with a separate control group.

We found significant decreases in COVID-19 case growth rate, death growth rate, and reproduction number during and in the later stage of the lockdown.

Title:- The COVID-19 Pandemic: A Pandemic of Lockdown Loneliness and the Role of Digital Technology

Link:- <https://doi.org/10.2196/22287>

Conclusion from paper:-

Digital technology has undoubtedly become critical for reducing and preventing social, physical, and psychological risks during the COVID-19 pandemic and addressing the short- and long-term impacts of social isolation and lockdown loneliness [18]. Nonetheless, most people affected by social isolation and lockdown loneliness during the pandemic might not feel lonely yet because these effects may take some time to show up [56]. It is therefore imperative that digital technology should not only provide tools to improve social connectedness and help in reducing lockdown loneliness but also enable people at risk of loneliness to take measures to avoid social isolation during the COVID-19 pandemic and in its aftermath. However, access to and costs and knowledge of digital technology tools are among the key issues that need urgent attention.

Finally, tackling lockdown loneliness will require the active involvement of all key stakeholders that use these digital technology tools.

Conclusion from model:-

Digital technology has undoubtedly become critical for reducing and preventing social, physical, and psychological risks during the COVID-19 pandemic and addressing the short- and long-term impacts of social isolation and lockdown loneliness [18].

Title:- The Impact of COVID-19 Lockdowns on Mental Health Patient Populations: Evidence from Medical Claims Data

Link:- <https://doi.org/10.1101/2021.05.26.21257598>

Conclusion from paper:-

Early in March 2020, non-pharmaceutical interventions, such as social distancing policies, were imposed around the world to contain the spread of COVID-19 and proved to reduce the number of COVID-19 cases and fatalities [23, 24, 12]. Mitigation policies come with both costs and benefits, which may be further analyzed to help determine the optimal time to release or stop a policy intervention [33]. Prior research showed significant mental health degradation associated with COVID-19 pandemic [17, 27, 2, 7, 19], however no research investigated the causal relation between COVID-19 mitigation policies and the usage of mental health resources. Considering the effects on the usage of mental health resources can further reflect the economic and health costs brought by the pandemic interventions. In our study, using large-scale medical claims data, we estimated the effects of lockdowns on the usage of mental health facilities and the prevalence of mental health issues at the state- and county-levels in the United States. Our findings demonstrate a statistically significant causal effect of lockdown measures (stay-at-home and school closure orders) on the usage of mental health facilities represented by increasing number of issued medical claims for mental health appointments during COVID-19 pandemic. Also, ED visits were statistically significant and positive in locked-down regions which reflect the increase of emergent mental help-seeking due to the COVID-19 lockdowns. Results further emphasize the cost brought by extra months of lockdowns, in which effect sizes keep increasing through the end of 2020 in both mental health visits and ED visits. Some sub-population groups were exposed to larger deterioration effect than other groups, such as female and adolescent groups. Given the various intertwined events and causes during the COVID-19 pandemic, our analysis is limited by several factors. First, it is important to point out that the adoption of lockdowns across states did not happen at random. Differences of shutdown orders' timings and adoption across regions were associated with the differences in COVID-19 confirmed cases and fatality rates across those regions [29, 35] and the differences in their health systems capacity [28]. Also, there

exist other political, economic, and institutional factors that affect the adoption of COVID-19 measures and their strictness level across countries [20]. Even though the lockdown timing may be affected by regional factors related to the virus, such as number of cases or institutional factors, however, there is no reason to believe that lockdown timing was affected the prevalence of mental health in regions. Given that, we have also encountered regional fixed effect in our model to adjust for regional differences. Second, though mental illnesses have a negative economic impact [43], the opposite is true as well, in which economic disadvantage may lead to a greater mental illness [32]. During COVID-19, there have been negative consequences on individuals in different industry sectors who were more likely to lose their jobs due to the lockdown measures [18] with significant employment loss in occupations that require interpersonal contact [39]. Therefore, the loss of employment due to shutdowns may have a confounding effect on increased mental health issues. In addition, the medical claims used in this study do not cover Medicare and Medicaid health insurance programs which creates a limitation on our data. Medicare covers most aged and disabled population across the US, while Medicaid covers a wider range of population including low-income beneficiaries covering 30% of US population [31]. This limitation would impact the representativeness of results, since our data misses some population groups in the US. Despite the mentioned limitations, our results provide important policy implications from economic and social impacts. There is a notable mental health cost brought by non-pharmaceutical interventions, especially interventions that are extended to longer duration. Our results suggest that there should be considerations to the mental health cost through ensuring mental health treatment capacity. Furthermore, we showed that number of patients had dropped right after lockdowns and then progressively increased in June and July 2020, supporting the findings of [50, 41]. This suggests that people with mental health afflictions did not have the ability to seek immediate care during restrictive lockdowns. Findings suggest that policy interventions should be accompanied with strategies that facilitate mental health treatment despite restrictive lockdowns, in order to avoid the exacerbated effect of delayed treatment.

Conclusion from model:-

Our findings demonstrate a statistically significant causal effect of lockdown measures (stay-at-home and school closure orders) on the usage of mental health facilities represented by increasing number of issued medical claims for mental health appointments during COVID-19 pandemic.

Even though the lockdown timing may be affected by regional factors related to the virus, such as number of cases or institutional factors, however, there is no reason to believe that lockdown timing was affected the prevalence of mental health in regions. In our study, using large-scale medical claims data, we estimated the effects of lockdowns on the usage of mental health

facilities and the prevalence of mental health issues at the state- and county-levels in the United States. Findings suggest that policy interventions should be accompanied with strategies that facilitate mental health treatment despite restrictive lockdowns, in order to avoid the exacerbated effect of delayed treatment.

During COVID-19, there have been negative consequences on individuals in different industry sectors who were more likely to lose their jobs due to the lockdown measures [18] with significant employment loss in occupations that require inter personal contact [39].

Title:- Changes in socio cultural attitudes towards appearance, body image, eating attitudes and behaviors, physical activity, and quality of life in students before and during COVID-19 lockdown

Link:- <https://doi.org/10.1016/j.appet.2021.105452>

Conclusion from paper:-

Overall, our findings demonstrate that COVID-19–related lockdowns have had a unique impact on the sociocultural attitudes towards appearance, health-related lifestyles and quality of life of students of either gender. In men, stay at home situations have significantly increased the internalization of thin/low fat beauty ideals and time spent browsing the internet for self-pleasure, decreased physical activity levels, lowered self-rated health and decreased unhealthy nutritional habits. No changes were observed in the BMI, body image, self-esteem or quality of life of males. In women, the lockdown increased the time spent browsing the internet for self-pleasure, the internalization of thin/ low fat beauty ideals, BMI, while also decreasing unhealthy nutritional habits and increasing sleep duration. No changes were observed in the physical activity levels of females. For women, the lockdown significantly increased satisfaction with body appearance, as well as general and psychological quality of life. These findings suggest that the majority of students cope with lockdown–related situation well. However, according to the TIM drastic increases in the internalization of stereotyped thin/low body fat ideals might trigger body image concerns and increase disordered eating after the lockdown. Helping both male and female students cope with the outcomes of lockdown-specific interventions may be beneficial. Specific interventions that help to decrease internalization of stereotyped beauty ideals and aim to promote positive body image and physical activity in students of universities during and after the lockdown might be beneficial.

Conclusion from model:-

In men, stay at home situations have significantly increased the internalization of thin/low fat beauty ideals and time spent browsing the internet for self-pleasure, decreased physical activity

levels, lowered self-rated health and decreased unhealthy nutritional habits. In women, the lockdown increased the time spent browsing the internet for self-pleasure, the internalization of thin/low fat beauty ideals, BMI, while also decreasing unhealthy nutritional habits and increasing sleep duration. **Conclusions**

Overall, our findings demonstrate that COVID-19–related lockdowns have had a unique impact on the sociocultural attitudes towards appearance, health-related lifestyles and quality of life of students of either gender. Specific interventions that help to decrease internalization of stereotyped beauty ideals and aim to promote positive body image and physical activity in students of universities during and after the lockdown might be beneficial. However, according to the TIM drastic increases in the internalization of stereotyped thin/low body fat ideals might trigger body image concerns and increase disordered eating after the lockdown.

Title:- Effects of the COVID-19 pandemic and lockdown on alcohol use disorders and complications

Link:- <https://doi.org/10.1097/YCO.0000000000000720>

Conclusion from paper:-

In this review to understand the effects of alcohol during the COVID-19 pandemic, changes in the pattern of use of alcohol and resultant effects are evident. Immediate effects have been an increase in alcohol-related emergencies including alcohol withdrawal, withdrawal-related suicides, methanol toxicity and alcohol-related motor vehicle accidents. Although the results on changes in alcohol use patterns during lockdown are mixed, there have been reports of binge/heavy drinking during lockdown and relapses post lockdown. Multiple psychological, social, biological, economic and policy related factors influence changes in drinking. A study from Switzerland shows that on an average, a person would lose 0.205 Years of Lost Life (YLL) due to psychological consequences of COVID-19, including alcohol use. This loss would be borne by 2.1% of the population who in turn would suffer an average of 9.79 YLL [116]. Hence, steps to optimize resources and to mitigate suffering in the most affected populations is necessary. Among patients with alcohol-related liver disease, caution is warranted related to use of medications, and outcomes appear to be worse. It has also been observed that alcohol increased the risk of COVID-19 infection. Services for patients with alcohol use disorders have been adversely affected across the globe. It is thus important to focus and train healthcare workers like nursing health professionals to deliver addiction related services [117]. Another important way to reduce the treatment gap is to harness technology [57]. One such example is e-consult for people with SUD during the pandemic and training support to healthcare workers in distant places to manage with SUD [118]. Telehealth, group meetings and online consultations

can be some ways to handle the increased demand during and after pandemic [119]. Although some nations banned alcohol sales completely others declared it as an essential product, resulting in varied problems across countries, including unintended messaging that alcohol is ‘essential’ [120]. There is a need for a rethinking about policy changes like online alcohol delivery, which can be difficult to roll back [121]. Evidence based restriction of alcohol pricing, availability and marketing are required for the future [122]. Governments should refrain from abrupt and knee-jerk alcohol policy changes (either a sudden ‘ban’, ‘online sale of alcohol’, declaring alcohol as ‘essential’) and instead adopt evidence-based decision making. Adequate information to the public in the event of anticipated limitations of access and information on treatment services should be provided on priority. Further, governments should adopt adequate measures to protect vulnerable populations. Importantly, post hoc research on impact of such decisions need to be undertaken. Among vulnerable groups like health professionals, elderly, patients diagnosed with cancer, alcohol has added to the burden of the problem. As there are some countries in the second wave at the time of this review, we need newer protocols and cohorts to study the long-term effects on mental health and addiction of different populations [123]. Finally, adapting to the current situation and preparedness to handle the repercussions due to pandemic is important. It is important to focus on the preventive dimension and early intervention. Survivor guilt, PTSD among survivors may put people at risk to addiction [124]. Evidence based policy changes, improving access to treatment for alcohol use disorders, liaison services, evidence-based prevention, and prioritizing care of vulnerable population are urgently required [125]. Preparedness plans to handle such emergencies in future are also required.

Conclusion from model:-

Governments should refrain from abrupt and knee-jerk alcohol policy changes (either a sudden ‘ban’, ‘online sale of alcohol’, declaring alcohol as ‘essential’) and instead adopt evidence-based decision making. Evidence based policy changes, improving access to treatment for alcohol use disorders, liaison services, evidence-based prevention, and prioritising care of vulnerable population are urgently required [125]. Although some nations banned alcohol sales completely others declared it as an essential product, resulting in varied problems across countries, including unintended messaging that alcohol is ‘essential’

Immediate effects have been an increase in alcohol-related emergencies including alcohol withdrawal, withdrawal-related suicides, methanol toxicity and alcohol-related motor vehicle accidents.

In this review to understand the effects of alcohol during the COVID-19 pandemic, changes in the pattern of use of alcohol and resultant effects are evident. As there are some countries in the

second wave at the time of this review, we need newer protocols and cohorts to study the long-term effects on mental health and addiction of different populations [123].

Although the results on changes in alcohol use patterns during lockdown are mixed, there have been reports of binge/heavy drinking during lockdown and relapses post lockdown.

Discussion

Generally speaking, the summaries produced by the software represented good abstractions of the main points described in the articles. The software did have trouble with organization of information for the larger bodies of text that it tried to summarize. Even so, software such as this can serve as a powerful tool to help people sift through large quantities of information to distill out the key takeaways. This can prove a valuable resource for both laypeople and professionals.

In spite of the success of the present software, there are enhancements that could make the software even more useful. As is often the case, research and other papers present different or even conflicting findings. In the field of scientific publication, these differences and also similarities are aggregated and summarized in review papers. An obvious enhancement to the present software is to enable it to perform similar meta-analyses. In this enhancement, the software could separate studies by their findings and report studies that report consistent findings and those that disagree with those findings. The next step would be to look at differences between the studies to identify potential variables that can account for differences. This could be a valuable source of potential research questions to investigate.

The present software also requires its user to retrieve the scientific articles and then feed them into the software. This can be a time-intensive process. Another enhancement is to integrate the current technology with our automated search and retrieval technology that we reported in our previous papers (Boina et al., 2021a, b, c, d). This would make the tool even more powerful as it would then accept a topic from a user, perform the Internet search and retrieval of the technical articles and then summarize their main points. By combining these technologies and enhancing them with the capabilities described above, we can create a next generation of search engines that do more than just retrieve information but also provide its users with the key takeaways from that information.

References

Boina, N., Agarwal, J., Agarwal, T., Leddo, J. et al. (2021a). A Novel Meta-Machine Learning Approach to Diagnose Stress from Environmental Factors Using Automated Knowledge Graphs, *International Journal of Social Science and Economic Research*, 6(12), 4961-4970.

Boina, N., Agarwal, J., Agarwal, T., Leddo, J. et al. (2021b). A Novel Meta-Machine Learning Approach to Diagnose Stress from Individual Factors Using a Self-retrieved Dataset and then Provide Directed Treatment, *International Journal of Social Science and Economic Research*, 6(12), 4933-4944.

Boina, N. et al. (2021c). A Novel Meta-Machine Learning Platform Able to Autonomously Learn How to Diagnose Acne and Jaundice. *International Journal of Social Science and Economic Research*, 6(10), 4151-4158.

Boina, N. et al. (2021d). A Novel Meta-Machine Learning Platform Able to Autonomously Learn How to Diagnose Autism, Breast Cancer, Melanoma Mole Cancer and Pink Eye. *International Journal of Social Science and Economic Research*, 6(10), 4159-4171.

Ecmekecioglu, C. (2020). Nutrition and Longevity – From mechanisms to uncertainties. *Critical Reviews in Food and Science Nutrition*, 60(18), 3063-3082.

Longo, V.D., Di Tano, M., Mattson, M.P. & Guidi, M. P. (2021). Intermittent and periodic fasting, longevity and disease. *Nature Aging*, 1, 47-59.

Roberts, M.N., Wallace, M.A., Tomilov, A.A., Cortopassi, G.A., Ramsey, J.J.& Lopez-Dominguz, J.A. (2017). A Ketogenic Diet Extends Longevity and Healthspan in Adult Mice. *Cell Metabolism*, 26, 539-546.

Trichopoulou, A. & Vasilopoulou, E. (2000). Mediterranean Diet and Longevity. *British Journal of Nutrition*, Dec;84 Suppl 2:S205-9. doi: 10.1079/096582197388554. PMID: 11242471.