**Appendix A: Supplemental Material**

Related to NLP Project Submission from David Kurniadi, Rene Lizarraga, Xin Gu

**Q1:** Range of incubation periods for the disease in humans (and how this varies across age and health status) and how long individuals are contagious, even after recovery.

**GloVe-Cosine**

'In addition, the asymptomatic incubation period for individuals infected with 2019-nCov was estimated to range from 1 to 14 days (most likely 3e10 days), longer than that of SARS-CoV.'

'The incubation period was determined to be 7 days in January, and it was recently estimated to be 3 days, which means that 2019-nCoV infected people become more likely to develop symptoms.',

'Even SARS-CoV infection, which had a dramatic effect on adults, was mainly associated with relatively mild disease in almost all patients <12 years of age (28, 34).'

'Individuals become symptomatic after an average incubation period of 10 days (range 2-21 days) [12], and infectiousness is increased during the later stages of disease [13].'

'Usually, COVID-19 has an incubation period of 2-7 days 2 with no obvious 41 symptoms, during which time the virus can spread from infected to uninfected individuals.',

**LDA-GloVe-Cosine**

A detailed analysis of one of the early COVID-19 clusters by Chan and colleagues 19 revealed symptomatic infections in five adult members of the same household, while a child in the same household aged 10 years was infected but remained asymptomatic, potentially indicating biological differences in the risk of clinical disease driven by age.

Our line list comprised 507 patients reported from Jan 13, to Jan 31, 2020, including 364 (72%) from mainland China and 143 (28%) from outside of China (table) The age distribution of COVID-19 cases was skewed towards older age groups with a median age of 45 years (IQR 33-56) for patients who were alive or who had an unknown outcome at the time of reporting (figure 1).

**Q2:** Prevalence of asymptomatic shedding and transmission (e.g., particularly children).

**GloVe-Cosine**

'Patients with cough had significantly higher viral loads than those who did not have these signs/symptoms, irrespective of age or virus type (Fig 2B-2E-2H).'

'The only significant differences were mean age (which was significantly higher in the children with a single viral infection than in those with coinfections) (P < 0AE001), and the radiographic finding of alveolar pneumonia, which was significantly more frequent in the children with co-infections (P < 0AE05).',

**LDA-GloVe-Cosine**

After the stage, likely after mid-January, 2020, the virus further spread to the family via infected adults to cause intrafamilial transmission, especially transmission to the elderly and children, who are vulnerable to the infection.

1B , during the emerging stage of the SARS-CoV-2 outbreak, the infection was disseminated by person-to-person transmission in the community almost exclusively among adults.

First, asymptomatic cases were diagnosed based on positive viral nucleic acid test results, but without any COVID-19 symptoms, such as fever, gastrointestinal, or respiratory symptoms, and no significant abnormalities on chest radiograph 7, 8 However, the transmission of COVID-19 through asymptomatic carriers via person-to-person contact was observed in many reports.

It is essential to know the incubation period, the time elapsing between the moment of exposure to an infectious agent and the appearance of signs and Furthermore, patients with pneumonia were older, with a higher prevalence of smoking history, more underlying diseases, and were more likely to have fever, myalgia/fatigue, dyspnea, headache, and nausea/vomiting compared to patients with ARD (all p < .05) ( Table 3 ).

In children, however, white blood cell count and absolute lymphocyte count were mostly normal, and no lymphocyte depletion occurred, suggesting less immune dysfunction after the SARS-CoV-2 infection.

Usually they recover within 1-2 weeks after the onset of the disease.

**Q3:** Seasonality of transmission.

#Since Covid19 is only started a couple month ago, not enough research about seasonality transmission. Something we know is that Covid19 and SARS are similar, they may have similar seasonality transmission pattern.

**GloVe-Cosine:**

None

**LDA-GloVe-Cosine:**

Both outbreaks happened in the winter, when the two provinces have similar climate patterns suitable for virus survival and spread.

**Q4:** Physical science of the coronavirus (e.g., charge distribution, adhesion to hydrophilic/phobic surfaces, environmental survival to inform decontamination efforts for affected areas and provide information about viral shedding).

# no closely related result sentences found

**GloVe-Cosine**

'Based on a lack of epidemiological data that would provide evidence of a benefit for the patient from surface disinfection (e.g., from a significant reduction of nosocomial infection rates), some scientists postulate that cleaning of surfaces with non-antimicrobial detergents is generally sufficient [1]

**LDA-GloVe-Cosine**

None

**Q5:** Persistence and stability on a multitude of substrates and sources (e.g., nasal discharge, sputum, urine, fecal matter, blood).

# LDA

## Article

Aerodynamic Characteristics and RNA Concentration of SARS-CoV-2 Aerosol in Wuhan Hospitals during COVID-19 Outbreak

## Discussion (<https://www.biorxiv.org/content/10.1101/2020.03.08.982637v1.full>)

“This study also recorded an elevated airborne SARS-CoV-2 concentration inside the patient mobile toilet of Fangcang Hospital. This may come from either the patient's breath or the aerosolization of the virus-laden aerosol from patient’s faeces or urine during use. Ong et al. has found the wipe samples from room surfaces of toilets used by SARS-CoV-2 patients tested positive.[11](https://www.biorxiv.org/content/10.1101/2020.03.08.982637v1.full#ref-11) Our finding has confirmed the aerosol transmission as an important pathway for surface contamination. We call for extra care and attention on the proper design, use and disinfection of the toilets in hospitals and in communities to minimize the potential source of the virus-laden aerosol.”

# GloVe-Cosine

‘A liquid extract of each fecal sample (1:5 dilution) was used for NMH analysis and fecal NMH concentrations were back-calculated for the wet weight of the fecal samples and expressed in ng/g feces.',

 'The presence of surrounding organic material (e.g., blood, saliva, mucus, etc.)',

 'Viral RNAs could be found in nasal discharge, sputum, and sometimes blood or feces.'

# LDA-GloVe-Cosine

“Our findings add support to a hypothesis that virus-laden aerosol deposition may play a role in surface contamination and subsequent contact by susceptible people resulting in human infection.”

[Aerodynamic Characteristics and RNA Concentration of SARS-CoV-2 Aerosol in Wuhan Hospitals during COVID-19 Outbreak](http://doi.org/10.1101/2020.03.08.982637)

**Q6:** Persistence of virus on surfaces of different materials (e,g., copper, stainless steel, plastic).

# LDA

## Article

Potential role of inanimate surfaces for the spread of coronaviruses and their inactivation with disinfectant agents

## Summary (<https://www.sciencedirect.com/science/article/pii/S2590088920300081?via%3Dihub>)

“In a recent review on the persistence of human and veterinary coronaviruses on inanimate surfaces it was shown that human coronaviruses such as Severe Acute Respiratory Syndrome (SARS) coronavirus, Middle East Respiratory Syndrome (MERS) coronavirus or endemic human coronaviruses (HCoV) can persist on inanimate surfaces like metal, glass or plastic for up to 9 days.”

# GloVe-Cosine

'Non-animal food-contact swabs were analyzed by surface type (metal garage, metal tabletop, concrete floor, and rubber boot bottoms worn during the experiment).',

 'It can survive on plastic surfaces, stainless steel, glass slides, and paper files.',

 'Thus, the material itself (vinyl chloride, aluminum, plastic, stainless steel) had no effect on the diminishing bacterial numbers under dry conditions.'

# LDA-GloVe-Cosine

“Data on the transmissibility of coronaviruses from contaminated surfaces to hands were not found.”

[Potential role of inanimate surfaces for the spread of coronaviruses and their inactivation with disinfectant agents](http://doi.org/10.1016/j.infpip.2020.100044)

**Q7:** Natural history of the virus and shedding of it from an infected person

# LDA

## Article

Modelling the epidemic trend of the 2019 novel coronavirus outbreak in China

## Abstract (<https://www.biorxiv.org/content/10.1101/2020.01.23.916726v1.full>)

“After ruling out possible influenza and other coronaviruses by laboratory testing, the Chinese authorities isolated a new type of coronavirus (novel coronavirus, nCoV) on 7th January 2020, which was then named 2019-nCoV by WHO on 12th January [[1](https://www.biorxiv.org/content/10.1101/2020.01.23.916726v1.full#ref-1)]. As of 22nd January, 571 confirmed cases (including 15 medical staff) and 17 deaths have been reported in China, and 6 cases confirmed overseas [[2](https://www.biorxiv.org/content/10.1101/2020.01.23.916726v1.full#ref-2)]. On 20th January 2020. two local infections in the Chinese province of Guangdong with no direct visit to Wuhan were the first confirmed human-to-human transmission cases [[3](https://www.biorxiv.org/content/10.1101/2020.01.23.916726v1.full#ref-3)].”

# GloVe-Cosine

'Initially, the 2019-CoV outbreak was reported as limited person-to-person transmission and a contaminated source from infected or sick wild animals in the wet market may have been the common origin.'

# LDA-GloVe-Cosine

“Before viruses in wildlife make a jump to infect human beings, they usually accumulate a series of mutations in their viral genomes [42] and invade human beings as a result of human occupation of their normal ecosystem, as exemplified with a story of initial human infection by HIV carried by chimpanzees in rainforests of West Africa [43, 44].”

[Potential Factors Influencing Repeated SARS Outbreaks in China](http://doi.org/10.3390/ijerph17051633)

**Q8:** Implementation of diagnostics and products to improve clinical processes

# LDA

## Article

In silico approach to accelerate the development of mass spectrometry-based proteomics methods for detection of viral proteins: Application to COVID-19

## Summary (<https://www.biorxiv.org/content/10.1101/2020.03.08.980383v2>)

“We describe a method for rapid in silico selection of diagnostic peptides from newly described viral pathogens and applied this approach to SARS-CoV-2/COVID-19. This approach is multi-tiered, beginning with compiling the theoretical protein sequences from genomic derived data. In the case of SARS-CoV-2 we begin with 496 peptides that would be produced by proteolytic digestion of the viral proteins.“

# GloVe-Cosine

'There is a need for availability of data and information on humanvector-pathogen-ecosystem interfaces, drugs and vaccines development as well as diagnostics techniques and tools from preclinical to clinical levels.',

 'Improving the optimization and implementation of protocols suitable for clinical samples will no doubt improve microbial diagnosis in clinical practice.',

 'These techniques are expected to greatly accelerate the identification of specific phage binders, facilitating mAb development for use in research, clinical diagnostics, and pharmaceuticals for the treatment of human disease.'

# LDA-GloVe-Cosine

“The subfractionation of existing heparin preparations against anticoagulant activities (with proven low-toxicity profiles, good bioavailability and industrial-scale manufacturing) for off-label pathologies, provides an attractive strategy for quickly and effectively responding to COVID-19 and for the development of next generation heparin-based therapeutics.”

[Mycroft-West et al. (2020) Running title: SARS-CoV-2 surface S1 Receptor Binding Domain binds heparin The 2019 coronavirus (SARS-CoV-2) surface protein (Spike) S1 Receptor Binding Domain undergoes conformational change upon heparin binding](http://doi.org/10.1101/2020.02.29.971093)

**Q9:** Disease models, including animal models for infection, disease and transmission

# LDA

## Article

Clinical Medicine - Optimization Method for Forecasting Confirmed Cases of COVID-19 in China

## Introduction (<https://www.mdpi.com/2077-0383/9/3/674/htm>)

“A large family of viruses, called coronaviruses, are severe pathogens for human beings, which infect respiratory, hepatic, gastrointestinal, and neurologic diseases. They are distributed among humans, birds, livestock, mice, bats, and other wild animals [[**1**](https://www.mdpi.com/2077-0383/9/3/674/htm#B1-jcm-09-00674),[**2**](https://www.mdpi.com/2077-0383/9/3/674/htm#B2-jcm-09-00674),[**3**](https://www.mdpi.com/2077-0383/9/3/674/htm#B3-jcm-09-00674)]. The outbreaks of two previous coronaviruses, SARS-CoV and MERS-CoV in 2003 and 2012, respectively, have approved the transmission from animal to animal, and human to human [[**4**](https://www.mdpi.com/2077-0383/9/3/674/htm#B4-jcm-09-00674)].”

# GloVe-Cosine

'The MA15 virus will enhance the use of the mouse model for SARS because infection with this virus in mice reproduces many aspects of severe human disease, including morbidity, mortality, and pulmonary pathology.',

 'Ferrets are commonly used as experimental models of infection for a variety of respiratory viruses due to their susceptibility to these viruses and the close resemblance of the pathological features to those found in human infections [11, 12] , including the development of severe respiratory and neurological disease during henipavirus infection [13] [14] [15] .'

# LDA-GloVe-Cosine

“It is now hypothesized that one of the reservoir coronavirus species in bats crossed the species barrier to an intermediate mammal host (presumed to be a masked civet) sold at the wet market at the epicenter of the current epidemic, with subsequent mutation and transmission to humans, initiating the present epidemic of COVID-19.

”

Six weeks into the 2019 coronavirus disease (COVID-19) outbreak- it is time to consider strategies to impede the emergence of new zoonotic infections (<https://journals.lww.com/cmj/Citation/9000/Six_weeks_into_the_2019_coronavirus_disease.99378.aspx>)

**Q10:** Tools and studies to monitor phenotypic change and potential adaptation of the virus

**GloVe-Cosine**

'Due to the sequence similarity between the PKs (figure 1C), it is speculated that they 88 were formed by duplication events during viral replication, probably involving recombination.',

**LDA-GloVe-Cosine**

'In this study, we used TWIRLS, a machine-based approach to collect, summarize, and analyze about 15,000 biomedical articles related to coronavirus, with the aim to elucidate the mechanisms underlying coronavirus-induced host pathological changes.',

**Q11:** Immune response and immunity

**GloVe-Cosine**

None.

**LDA-GloVe-Cosine**

'We saw what appears to be an innate immune response at the 10,000 PFU EBOV exposure level.',

'It has been suggested that EBOV can mediate an innate immunity response through stimulation of TLR-4 [28] .',

'Hence, the S2 subunit may serve as an important antigen for inducing 2 0 2 both humoral as well as cell-mediated immunity against SARS-CoV and SARS-CoV-2.',

'mouse using an adjuvanted microsphere peptide vaccine formulation containing NP44-52 is enough to confer immunity in mice.',

**Q12:** Effectiveness of movement control strategies to prevent secondary transmission in health care and community settings

**GloVe-Cosine**

None

**LDA-GloVe-Cosine**

'This type of transmission would make effective contact tracing challenging, and good respiratory and hand hygiene would be crucial to reduce this route of transmission, coupled with environmental decontamination in health-care settings.',

'Future studies based on larger samples of patients with COVID-19 could explore in more detail the transmission dynamics of the outbreak in different locations, the effectiveness of interventions, and the demographic factors driving transmission.',

'Closure of certain routes, targeted airport screening, risk communication, public awareness and targeted training and vigilance of health workers associated with the portals of entry of visitors to their countries will help mitigate the force of further spread of 2019-nCoV.',

'The model could be modified to include some transmission after isolation (such as in hospitals), which would decrease the probability of achieving control.',

**Q13:** Effectiveness of personal protective equipment (PPE) and its usefulness to reduce risk of transmission in health care and community settings, mask, google, gloves

**GloVe-Cosine**

None

**LDA-GloVe-Cosine**

None

**Q14:** Role of the environment in transmission

**GloVe-Cosine**

None

**LDA-GloVe-Cosine**

'There is an urgent need for the implementation of multidisciplinary One Health to address the current complex health challenges at the human-animal-environment interface [42, 43] One Health approaches in China have recently been described [49] [50] [51] [52] .',

'Multidisciplinary research in biomedical, social, and environmental sciences is required to achieve a deeper understanding of disease transmission and develop more effective systems for emergency response.',

'The spread of infectious diseases is affected not only by the biological characteristics of the pathogen but also by various other factors such as politics, culture, economy, and the environment.',