

# Co-distribution of Light at Night (LAN) and COVID-19 Incidence in the United States

**Yidan Meng**

Yale University School of Public Health

**Vincent Zhu**

Yale University School of Public Health

**Yong Zhu** (✉ [yong.zhu@yale.edu](mailto:yong.zhu@yale.edu))

Yale University School of Public Health <https://orcid.org/0000-0002-6957-1114>

---

## Research article

**Keywords:** Light at night (LAN), COVID-19

**DOI:** <https://doi.org/10.21203/rs.3.rs-195410/v2>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

# Abstract

**Background:** Light at night (LAN) as a circadian disruption factor may affect the human immune system and consequently increase an individual's susceptibility or the severity of infectious diseases, such as COVID-19. COVID-19 infections spread differently in each state in the United States (US). The current analysis aimed to test whether there is an association between LAN and COVID-19 cases in 4 selected US states: Connecticut, New York, California, and Texas.

**Methods:** We analyzed clustering patterns of COVID-19 cases in ArcMap and performed a multiple linear regression model using data of LAN and COVID-19 incidence with adjustment for confounding variables including population density, percent below poverty, and racial factors.

**Results:** Hotspots of LAN and COVID-19 cases are located in large cities or metro-centers for all 4 states. LAN intensity is associated with cases/1k for overall and lockdown durations in New York and Connecticut ( $P < 0.001$ ), but not in Texas and California. The overall case rates is significantly associated with LAN in New York ( $P < 0.001$ ) and Connecticut ( $P < 0.001$ ).

**Conclusions:** We observed a significant positive correlation between LAN intensity and COVID-19 cases-rate/1k, suggesting that circadian disruption of ambient light may increase the COVID-19 infection rate possibly by affecting an individual's immune functions. Furthermore, differences in the demographic structure and lockdown policies in different states play an important role in COVID-19 infections.

## Background

Mammalian circadian rhythms, controlled by a neurological master clock located in the suprachiasmatic nucleus (SCN) and the peripheral clocks of somatic cells, regulate a number of biological and physiological processes including the human immune system [1, 2]. Because immune responses play a major role in fighting against virus infections [3], disrupted circadian rhythm may adversely influence immune functions and consequently increase virus infectivity and its ability to replicate inside hosts [4-6].

Circadian disruptions can be caused by sleep deprivation, night shift work, frequent air traveling, circadian gene alterations, and light at night (LAN) exposures [7-10]. Sleep deprivation has been associated with increased susceptibility to gut infection [11]. A higher incidence and severity of respiratory infections has been reported among night shift workers [12]. These findings support a significant relationship between disrupted circadian rhythms and an individual's increased vulnerability to infectious diseases and suggest that excess risk could also be observed among individuals with high LAN exposures for the infection of COVID-19 [2], a coronavirus causing the global pandemic in 2020.

There are also studies analyzing LAN exposure and various cancer types, such as breast cancer, prostate cancer, thyroid cancer, and non-Hodgkin Lymphoma [7, 9, 13, 14]. Both global and regional studies have shown that there is a significant association between intensity of light at night and breast cancer [7, 14-16]. These findings suggest that LAN as a circadian disruption can influence the immune system and hormone releases, and in turn affect an individual's susceptibility to infectious diseases as well.

Light at night comes from either ambient light or indoor artificial light exposures. Excessive exposure to LAN may generate light pollution that causes adverse effects on immune functions [17] and alters circadian gene functions in the SCN [13]. City-level LAN intensity can be measured by using the U.S. Defense Meteorological Satellite Program (DMSP). In the present analysis, we investigate whether exposure to LAN is associated with COVID-19 incidence in major cities in four selected US states: Connecticut, New York, Texas, and California, each of which represent different geological locations.

## Method

*Data sources* We obtained COVID-19 cases and testing data from local health departments. Specifically, we obtained COVID-19 data for Connecticut from the Connecticut State Department of Public Health (<https://data.ca.gov/dataset/covid-19-cases>), for New York State from the Open NY Program (<https://data.ny.gov/browse?tags=covid-19>), for Texas from Texas Health and Human Services (<https://dshs.texas.gov/coronavirus/additionaldata.aspx>), and for California from the California Open Data Portal (<https://data.ca.gov/dataset/covid-19-cases>). Data of COVID-19 cases were categorized in 3 groups: duration of overall period (to August 24<sup>th</sup>, 2020), period of lockdown, and period of reopening in each state according to local state policy on their government websites. In summary, we obtained 62 data points for New York State, 167 data points for Connecticut, 56 data points for California, and 254 data points for Texas, based on the data availability for either county or town level.

LAN intensity data was extracted from satellite images of nighttime light intensity created by NASA Earth Observatory [18]. We also collected demographic data, including factors of income and poverty, race/ethnicity, and population density. County level demographic data for California, Texas, and New York was obtained from the US Census (<https://data.census.gov>), and town level demographic data for Connecticut was obtained from local public health departments (<https://portal.ct.gov/DPH>). County and town level boundaries data was obtained from local transportation or planning departments (<http://gis.ny.gov/gisdata>; <https://data.ct.gov>; <https://gis.txdot.opendata.arcgis.com>; <https://gisdata-caltrans.opendata.arcgis.com>).

*Geographic information system (GIS) mapping* ArcMap (<https://desktop.arcgis.com/en/arcmap>) was used to generate visualized hotspot or density maps for COVID-19 case rate data and LAN data. The *Kernel Density* (KD) in ArcMap 10.8.1 was used to calculate density from neighborhood features to create a smooth raster layer from points or polylines. Search radius was calculated by spatial configuration and the total number of points in the dataset, and equal breaks were used for symbology for Connecticut, Texas, and California. Natural breaks were used for symbology for New York because the LAN level in New York City is much higher than other cities in New York and natural breaks can better present data with large differences of inherent groups. Because of the huge differences in LAN data of New York City compared to other cities in New York State, we specifically analyzed the spatial pattern of LAN and COVID-19 case rates and performed the Geographically Weighted Regression (GWR). The Spatial Autocorrelation (Global Moran's I) tool was used to test spatial patterns (clustered, dispersed, or random) of points for cases/1k

during lockdown, reopening, and overall, for New York State. GWR was used to understand regional variation of geo-data.

*Statistically analysis* To analyze the correlation between LAN intensities and COVID-19 case rates per 1,000 people, we built multiple linear regression models with variables of nonwhite rate, percent below poverty, and population density from the US Census, 2015: ACS 5-Years Estimates Subject Tables (<https://data.census.gov>). Multiple regression models were performed using SAS 9.4 and RStudio.

## Results

To represent the different regions of the U.S., we included four states—Connecticut, New York, California, and Texas—in the study. Generally, hotspots of LAN data were located in large cities or metro-centers for all four states tested, such as Hartford, CT, New York City, NY, Dallas, TX, and San Francisco, CA.

The maximum LAN intensity calculated by ArcMap tool was 254.68, 254.68, 235.14 and 190.30 in New York State, Connecticut, Texas, and California, respectively. The mean LAN intensity was 92.31 in Connecticut and around 40 in the other three states. The maximum COVID-19 case-rate was in Texas among all four states, which was around 67 cases/1k people, followed by California (around 47 cases/1k people), New York (around 43 cases/1k people), and lowest in Connecticut (around 27 cases/1k people) at the time of data collection.

For Connecticut and New York State, the hotspots of LAN intensity, COVID-19 case rates/1k during lockdown, and COVID-19 case rates/1k during the overall duration shared similar patterns that clustered around major cities in the state (**Fig. 1**), but were slightly different after reopening. Differently, in Texas and California, there were inconsistent patterns of hotspots for the four interested variables, but they share the trend that the hotspot-areas were similar for reopening and overall durations.

During the lockdown period, the hotspots of COVID-19 case rates shared similar patterns with LAN data in Connecticut and New York, but had different geo-patterns in California and Texas. During the reopening period, the locations of hotspots of COVID-19 case rates were very different from the LAN data map in New York, Texas, and California. The hotspot patterns of COVID-19 case rates during the overall period were similar to the LAN map in Connecticut and New York. On the contrary, the overall hotspots of COVID-19 case rates were similar to the reopening period in California and Texas.

An analysis of multivariable regression models also revealed similar patterns to the ones that geo-patterns showed. The cross comparison among four states showed that there were statistically significant correlations between LAN intensity and cases/1k for the overall and lockdown durations in New York and Connecticut ( $P < 0.001$ ). There was no statistically significant association between LAN intensity and cases/1k, for the overall, lockdown, or reopening durations, in Texas and California (**Table.1**). The overall case rates were significantly associated with LAN in New York State ( $p < 0.001$ ) and Connecticut ( $p < 0.001$ ), in which every 1 unit increase of LAN had a 15.6% increase in the overall case rate in New York, and a 3.7% increase in Connecticut. The results of the lockdown period were similar to those for the overall period in

New York( $p<0.001$ ) and Connecticut ( $p<0.001$ ). During the reopening period, there was a significant small positive association between case rates and LAN data in Connecticut ( $p<0.001$ ). Based on R-squared results, the state-specific regression models could explain more variations in New York ( $R^2 = 0.80, 0.78$  and  $0.40$ ) and Connecticut ( $R^2=0.58, 0.57$  and  $0.21$ ), compared to the data in California ( $R^2 = 0.22, 0.37, 0.23$ ) and Texas ( $R^2 = 0.13, 0.02, 0.12$ ), for the overall, lockdown and reopening durations.

## Discussion

Our findings demonstrate significant positive correlations between LAN intensity and COVID-19 case rates per 1000 in two of the four US states (Connecticut and New York) studied. These findings support the proposed hypothesis that high LAN exposures may disrupt circadian rhythms and consequently increase an individual's vulnerability to infectious diseases including COVID-19 [2].

The co-distribution of light at night (LAN) and COVID-19 incidence was more evident in New York State and Connecticut. The regression models we built can explain more variations of COVID-19 data in New York and Connecticut than in Texas and California. This difference might be due to different lockdown/reopening policies in each state. For example, California had a shorter lockdown period, and there were no formal Stay at Home orders in Texas, compared to New York which has similar total COVID-19 cases and population. The less strict policies may introduce more social factors into the analysis that we cannot determine easily at this point. We also found some inconsistent results of the effect of factors of social determinants, which might be due to the different social and demographic structures in each state. Moreover, the different COVID-19 testing policies and availability in different states might introduce more variations into our analysis.

Data points for all four states combined are relatively less likely to have a very solid statistical analysis, but in order to discuss different policy impacts, we chose to analyze states because of the heterogeneity of each state. The data used in the study is at either town or county levels, which aggregate individual data into large spatial area levels and may introduce ecological fallacy. To reduce the effect of ecological bias, we performed GWR models for New York. This approach improved the analysis and generated a very high R-squared value from the regression models (**Supplement Table 1**).

The current analysis focused on the COVID-19 test rate, and information on the severity of infected patients were not included due to its unavailability. Confounding factors considered in the analysis were percent below poverty, non-white rate, and population density. More confounders, such as employment, education level, underlying health conditions, and proximity to healthcare facilities should also be considered in future studies if they are available.

Observations from our study are consistent with findings from a recent study that shows melatonin usage is significantly associated with a 28% reduced likelihood of a positive laboratory test result for COVID-19 [19]. Melatonin production in the pineal gland is sensitive to light and it has shown that even exposures to low intensity light can suppress melatonin secretion [20]. LAN may reduce melatonin levels and consequently increase risk of COVID-19 infection.

Due to mental pressure, behavior, and daily routine changes during the pandemic, there are increasing concerns of sleep disturbances and circadian disruptions, especially for healthcare workers [5, 21, 22]. This provides more opportunities to analyze how circadian disruptions such as LAN correlate with hormone-related health outcomes and temporal immune dynamics [23].

## **Conclusion**

In summary, both LAN intensity and COVID-19 case rates are higher in major cities or metro-centers in all four states, due to the nature of cities of higher mobility, population density, etc. In the current study, we observed a significant positive correlation between LAN intensity and COVID-19 cases-rate/1000, which suggests that circadian disruption of ambient light may increase the COVID-19 infection rate possibly by affecting an individual's immune functions. Furthermore, differences in demographic structure and lockdown policies in each state play an important role in COVID-19 infections.

## **Declarations**

### **Ethical Approval and Consent to participate**

Not applicable

### **Consent for publication**

Not applicable

### **Availability of supporting data**

Not applicable

### **Competing interests**

The authors declare no potential conflicts of interest.

### **Funding**

YM was supported by internship funds from Yale University.

### **Authors' contributions**

All authors have read and approved the manuscript.

Acquisition of data: YM

Analysis and interpretation: YM, YZ

Writing, review, and/or revision of the manuscript: YM, VZ, YZ

Conception and design: YZ

## Acknowledgements

Not applicable

## Authors' information

Yidan Meng (yidan.meng@yale.edu), Vincent Zhu (vzhu2004@gmail.com), Yong Zhu (yong.zhu@yale.edu). Department of Environmental Health Sciences, Yale University School of Public Health, New Haven, Connecticut, 06520 USA

## References

1. Maiese K: **Moving to the Rhythm with Clock (Circadian) Genes, Autophagy, mTOR, and SIRT1 in Degenerative Disease and Cancer.** *Curr Neurovasc Res* 2017, **14**(3):299-304.
2. Silva FRD, Guerreiro RC, Andrade HA, Stieler E, Silva A, de Mello MT: **Does the compromised sleep and circadian disruption of night and shiftworkers make them highly vulnerable to 2019 coronavirus disease (COVID-19)?** *Chronobiol Int* 2020, **37**(5):607-617.
3. Rouse BT, Sehrawat S: **Immunity and immunopathology to viruses: what decides the outcome?** *Nat Rev Immunol* 2010, **10**(7):514-526.
4. Anderson G, Reiter RJ: **Melatonin: Roles in influenza, Covid-19, and other viral infections.** *Rev Med Virol* 2020, **30**(3):e2109-e2109.
5. Erren TC, Lewis P: **SARS-CoV-2/COVID-19 and physical distancing: risk for circadian rhythm dysregulation, advice to alleviate it, and natural experiment research opportunities.** *Chronobiol Int* 2020, **37**(7):1106-1109.
6. Meira E Cruz M, Miyazawa M, Gozal D: **Putative contributions of circadian clock and sleep in the context of SARS-CoV-2 infection.** *Eur Respir J* 2020, **55**(6):2001023.
7. Garcia-Saenz A, Sanchez de Miguel A, Espinosa A, Valentin A, Aragones N, Llorca J, Amiano P, Martin Sanchez V, Guevara M, Capelo R *et al.*: **Evaluating the Association between Artificial Light-at-Night Exposure and Breast and Prostate Cancer Risk in Spain (MCC-Spain Study).** *Environ Health Perspect* 2018, **126**(4):047011.
8. IARC: **Carcinogenicity of night shift work.** *Lancet Oncol* 2019;20(8):1058-1059 2019(1474-5488 (Electronic)).
9. Stevens RG, Zhu Y: **Electric light, particularly at night, disrupts human circadian rhythmicity: is that a problem?** *Philos Trans R Soc Lond B Biol Sci* 2015, **370**(1667).
10. Touitou Y, Reinberg A, Touitou D: **Association between light at night, melatonin secretion, sleep deprivation, and the internal clock: Health impacts and mechanisms of circadian disruption.** *Life Sci* 2017, **173**:94-106.

11. Anderson JR, Carroll I, Azcarate-Peril MA, Rochette AD, Heinberg LJ, Peat C, Steffen K, Manderino LM, Mitchell J, Gunstad J: **A preliminary examination of gut microbiota, sleep, and cognitive flexibility in healthy older adults.** *Sleep Med* 2017, **38**:104-107.
12. Archer SN, Oster H: **How sleep and wakefulness influence circadian rhythmicity: effects of insufficient and mistimed sleep on the animal and human transcriptome.** *J Sleep Res* 2015, **24**(5):476-493.
13. Stevens RG, Blask DE, Brainard GC, Hansen J, Lockley SW, Provencio I, Rea MS, Reinlib L: **Meeting report: the role of environmental lighting and circadian disruption in cancer and other diseases.** *Environ Health Perspect* 2007, **115**(9):1357-1362.
14. Zhong C, Franklin M, Wiemels J, McKean-Cowdin R, Chung NT, Benbow J, Wang SS, Lacey Jr JV, Longcore T: **Outdoor artificial light at night and risk of non-Hodgkin lymphoma among women in the California Teachers Study cohort.** *Cancer Epidemiology* 2020, **69**:101811.
15. Keshet-Sitton A, Or-Chen K, Huber E, Haim A: **Illuminating a Risk for Breast Cancer: A Preliminary Ecological Study on the Association Between Streetlight and Breast Cancer.** *Integr Cancer Ther* 2017, **16**(4):451-463.
16. Kloog I, Stevens RG, Haim A, Portnov BA: **Nighttime light level co-distributes with breast cancer incidence worldwide.** *Cancer Causes Control* 2010, **21**(12):2059-2068.
17. Nelson RJ: **Seasonal immune function and sickness responses.** *Trends Immunol* 2004, **25**(4):187-192.
18. **Global Map Downloads - 2016 Grayscale** [<https://earthobservatory.nasa.gov/features/NightLights>]
19. Zhou Y, Hou Y, Shen J, Mehra R, Kallianpur A, Culver DA, Gack MU, Farha S, Zein J, Comhair S *et al*: **A network medicine approach to investigation and population-based validation of disease manifestations and drug repurposing for COVID-19.** *PLoS Biol* 2020, **18**(11):e3000970.
20. Cajochen C, Munch M, Kobiacka S, Krauchi K, Steiner R, Oelhafen P, Orgul S, Wirz-Justice A: **High sensitivity of human melatonin, alertness, thermoregulation, and heart rate to short wavelength light.** *J Clin Endocrinol Metab* 2005, **90**(3):1311-1316.
21. Gao C, Scullin MK: **Sleep health early in the coronavirus disease 2019 (COVID-19) outbreak in the United States: integrating longitudinal, cross-sectional, and retrospective recall data.** *Sleep Med* 2020, **73**:1-10.
22. Morin CM, Carrier J, Bastien C, Godbout R, Canadian S, Circadian N: **Sleep and circadian rhythm in response to the COVID-19 pandemic.** *Can J Public Health* 2020, **111**(5):654-657.
23. Shirazi J, Donzanti MJ, Nelson KM, Zurakowski R, Fromen CA, Gleghorn JP: **Significant Unresolved Questions and Opportunities for Bioengineering in Understanding and Treating COVID-19 Disease Progression.** *Cell Mol Bioeng* 2020:1-26.

## Tables

**Table 1:** Summary table of Regression models of association between COVID case-rate and LAN intensity with covariants: nonwhite rate, percent below poverty, and population density for the overall, lockdown, and reopening periods for New York, Connecticut, California, and Texas.

Variables	N	Overall cases rate			Lockdown cases rate			Reopening cases rate		
		Beta	P-value	R <sup>2</sup>	Beta	P-value	R <sup>2</sup>	Beta	P-value	R <sup>2</sup>
<b>New York</b>	62	0.156	6.61E-06	0.80	0.148	1.34E-05	0.78	0.003	0.4072	0.40
<b>Connecticut</b>	167	0.037	6.10E-11	0.58	0.030	8.90E-11	0.57	0.006	0.0080	0.21
<b>Texas</b>	254	0.043	0.1602	0.13	0.001	0.8682	0.02	0.041	0.1558	0.12
<b>California</b>	56	0.110	0.1428	0.22	0.001	0.8960	0.37	0.110	0.1297	0.23

**Note:**

N: county level data points for states of NY, CA, TX; town level data points for CT;

Beta: beta-estimate for the regression model, representing the coefficient of models;

R<sup>2</sup>: measured the goodness of fitting for regression models, representing percent of data that is able to be explained by the model.