Agricultural Pesticide Use and Effects on Migrant Farmworker Health

MDI Scholars: Jasmine Jia, MS Data Science for Public Policy, 2025; Katharyn Loweth, MS Data Science for Public Policy, 2025; Vaishnavi Singh, MS Data Science for Public Policy, 2025;

Supervisor: Dr. Ioannis Ziogas
Funded By: The Lawyers' Committee for Civil Rights Under Law

Abstract

The research uses publicly available data sources to identify areas in the U.S. that have high numbers of non-U.S. citizen farmworkers and laborers and persons treated for non-Hodgkin's Lymphoma. This research will be used to conduct targeted outreach to identify individuals who have been diagnosed with specific form of cancer because of their potential exposure to glyphosate, but were excluded based on their citizenship status from previous settlements with pesticide manufacturers.

Introduction

Across the country, farms utilize pesticides like glyphosate to control weeds and grass. These pesticides are known to be detrimental to the health of those who are exposed to it. Over the last decade, farmers and landscapers have been awarded settlements in glyphosate litigation against the pesticide manufacturers. However, these settlements have not been able to support a particularly vulnerable part of the agricultural workforce: migrant and undocumented workers. These groups were previously excluded from U.S. settlements based on their citizenship status.

The Massive Data Institute collaborated with the Lawyers' Committee for Civil Rights Under Law in an effort to help identify those who were previously unable to benefit from settlements. Specifically, the MDI team worked to produce a data set that included data showing areas with both: (1) high numbers of non-citizen farmworkers and (2) persons treated for cancer. The Lawyer's Committee could then use these results to inform their targeted outreach efforts.

Acknowledgements

The MDI scholars would like to thank our supervisor Dr. Ioannis Ziogas for this support on this project, and the Lawyers' Committee for funding this work.

Methodology

Agricultural Pesticide Use:

Labor Intensity Weights: Each crop is assigned a labor intensity weight—high (3), medium (2), low (1), or tree-grown (0)—to reflect the relative amount of manual labor required. This categorization considers factors such as harvesting processes and potential exposure to herbicides like glyphosate.

Weighted Labor Acreage Calculation: The "weighted labor" metric for each crop within a county is calculated by multiplying a crop's planted acreage by its designated labor intensity weight. This method integrates both the scale of cultivation and the inherent labor demands of each crop type.

Migrant Farmworkers (H-2A Visas):

The estimated number of migrant farmworkers in each county and state is based on the Department of Labor (DOL)'s annual employer H-2A certification records from Fiscal Years (FYs) 2017-2021. We used the recorded worksite state and county, when available, and the number of workers certified for each employer. We used Google's Geoplaces API to ensure uniformity in records across years and to fill in missing information when appropriate. We aggregated the records together at the county/state level for each FY. We also used these estimates to develop a proxy measure for the number of undocumented farm workers as well.

Non-Hodgkin's Lymphoma Cancer Incidence:

Age-adjusted NHL incidence rates were collected from the NIH and CDC for the years 2017–2021, with population data from the US Census Bureau used to address gaps in county-level reporting. Population-adjusted rates were calculated using the formula: **Population-Adjusted NHL Rate**=(Average Annual Count/Population)×100,000

To categorize counties by risk level, thresholds were defined: high risk (above 314.25), moderate risk (116.88–314.25), and low risk (116.88 or below). Missing data due to suppression were estimated using population figures and average annual counts, ensuring a more complete dataset for analysis.

Calculation Metrics:

The labor demand and the migrant farmworker datasets were aggregated from 2017-2021 to match with the NHL Cancer Incidence. The identification of the top counties was based on taking the top 25% for NHL Cancer Incidence and labor demand, and then matching with the H2A dataset to identify gaps. Our hypothesis is that, in several counties, the formal H2A visa program did not adequately meet the labor demand, implying that undocumented immigrant workers are filling these gaps to ensure the agricultural operations remain viable.This approach highlighted counties with substantial agricultural labor needs, reliance on illegal labor, and elevated health risks for workers.

Findings

The top 10 counties are Fresno (CA), Yuma (AZ), Glenn (CA), Flagler (FL), Hillsborough (FL), Kern (CA), Moore (NC), Pierce (GA), Jefferson (ID), and St. Johns (FL). Geographically, several of these counties are located in states close to U.S. borders or coastal regions, such as California and Florida, which could correlate with higher labor migration and the presence of undocumented immigrants. These areas rely heavily on manual agricultural labor such as berry production because of the warm weather, and their proximity to borders may facilitate the movement of individuals seeking employment opportunities. For instance, Florida and California have substantial agricultural production, requiring intensive labor often supported by temporary foreign workers. Additionally, counties in states like Wisconsin and Georgia also exhibit high labor demands for agricultural production, reflecting the reliance on labor-intensive crop farming. The elevated NHL rates in these counties may be associated with exposure to hazardous conditions, such as pesticide use, commonly found in manual agricultural labor.

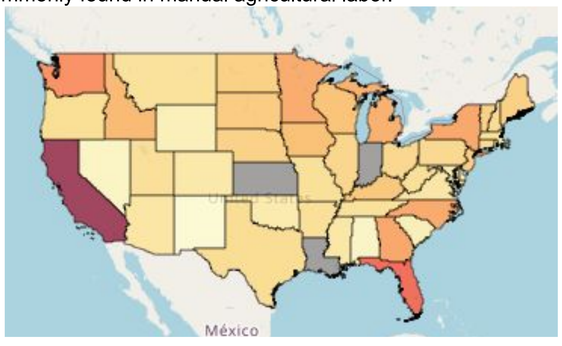

Figure 1 — U.S. States with High Level of Labor Demand, H-2A Workers, and NHL Rate

Figure 1 shows the distribution of U.S. states with varying levels of labor demand, H-2A workers, and NHL rates. States like California, Florida, Washington, and Arizona exhibit particularly high values across these metrics, indicating significant concentrations of labor needs and foreign worker utilization.

Limitations

The primary limitation the team experienced was lack of quality data at the county level. The incidence rate for non-hodgkin's lymphoma at a county level was not reported for many counties across the U.S, and was only available from 2017-2021. Additionally, there is little reliable public data on this undocumented community so we had to focus our analysis primarily on certified migrant workers.

Ethical Considerations

The purpose of our work is to help migrant and undocumented farmworkers who have developed serious health issues because of their pesticide exposure. However, while our intent is to create a dataset that identifies areas with likely high numbers of undocumented individuals and families that will be used to help this community, considering the current U.S. political climate this data could be used in ways that could harm this community instead. Thus while collecting data on this vulnerable group, we aim to protect these communities in our data as much as we can.

Conclusion

In summation, through collecting and combining agricultural, labor and health-focused datasets, we were able to identify counties in the U.S. that have above average presence of non-U.S. citizen farmworkers and non-Hodgkin's lymphoma cancer incidence.

References

1. Employment and Training Administration. (n.d.). *Performance Data*. U.S. Department of Labor. Retrieved from https://www.dol.gov/agencies/eta/foreign-labor/performance.
2. U.S. Center for Disease Control. (n.d.). *United States Cancer Statistics*. Retrieved from https://www.cdc.gov/united-states-cancer-statistics/dataviz/?CDC_AAref_Val=https://www.cdc.gov/cancer/uscs/dataviz/
3. U.S. Department of Agriculture, Farm Service Agency. (2024). *Crop Acreage Data*. Retrieved from https://www.fsa.usda.gov/tools/informational/freedom-information-act-foia/electronic-reading-room/frequently-requested/crop-acreage-data