Regional Disparities in Maternal and Infant Health: The Role of Gestational Diabetes and Hypertension Across U.S. States

Zaina Khan

Received October 19, 2024
Accepted January 24, 2025
Electronic access February 15, 2025

Gestational diabetes and hypertension significantly impact infant mortality rates, presenting notable public health challenges. This study investigates the correlations between these maternal health conditions and infant mortality rates across various U.S. states, aiming to elucidate their complex interactions and state-wise discrepancies. Comprehensive datasets encompassing state-level indicators of gestational diabetes, hypertension, and infant mortality were meticulously cleaned, standardized, and integrated for rigorous analysis. Employing correlation heatmaps, statistical tests, and advanced machine learning techniques, including K-means clustering and SVM classification, uncovered patterns and relationships within the data. Our findings reveal a weak positive correlation between gestational diabetes and infant mortality ($r = 0.12$, $p < 0.05$) and a stronger positive correlation between hypertension and infant mortality ($r = 0.23$, $p < 0.01$). States with higher prevalence levels of these maternal health conditions, such as Mississippi and Arkansas, exhibit notably higher infant mortality rates compared to states with lower prevalence levels, like Massachusetts. The regression model explained 28% of the variance in infant mortality rates ($R^2 = 0.28$), emphasizing the multifactorial nature of these outcomes. Cluster analysis revealed regional variations, pinpointing states in need of targeted interventions. These findings highlight the urgent need for tailored public health strategies to address maternal health disparities and improve infant survival rates. The study contributes to the growing evidence linking maternal health conditions to infant mortality and offers a robust framework for future public health policies. This research contributes to the growing body of evidence linking maternal health conditions to infant mortality, offering a comprehensive analytical framework for future studies and public health initiatives.

Introduction

Maternal health is a cornerstone of public health, closely linked to infant mortality rates, which serve as critical indicators of societal well-being and healthcare efficacy. Pre-pregnancy and gestational conditions, particularly hypertension and diabetes, significantly contribute to adverse maternal and neonatal outcomes[1]. These conditions not only pose immediate health risks but also have profound long-term implications for infant survival and development. Hypertension complicates approximately 6-8% of pregnancies globally and manifests as chronic hypertension predating pregnancy, gestational hypertension, or preeclampsia—a severe form involving hypertension and organ dysfunction[2]. This condition disrupts placental development and maternal vascular health, resulting in outcomes such as intrauterine growth restriction (IUGR), preterm birth, and elevated infant mortality rates[3]. Similarly, diabetes mellitus, both pre-existing and gestational, amplifies risks for congenital anomalies, macrosomia, and neonatal complications like hypoglycemia and respiratory distress[4]. This research aims to address these gaps by investigating the combined effects of hypertension and diabetes during pregnancy, focusing on their overlapping pathophysiological mechanisms and evaluating potential interventions to mitigate associated risks. By targeting these gaps, the study seeks to inform strategies for comprehensive prenatal care that optimize maternal and neonatal outcomes.

The interplay between maternal hypertension and diabetes presents a complex landscape for maternal-fetal health, with overlapping pathophysiological mechanisms and compounded risks. Women with coexisting hypertension and diabetes exhibit heightened susceptibility to adverse outcomes, including increased rates of preterm birth, fetal growth restriction, and perinatal mortality[5]. This dual burden underscores the urgency of comprehensive prenatal care and necessitates targeted interventions to mitigate risks and optimize outcomes for both mothers and infants. State-specific studies reveal substantial disparities in maternal health outcomes and the prevalence of gestational diabetes and hypertension. For instance, southern states such as Mississippi and Alabama report higher rates of maternal mortality and infant mortality, often correlated with limited access to prenatal care, higher rates of chronic conditions among pregnant individuals, and socioeconomic inequities[6]. For example, Mississippi recorded a maternal mortality rate of 33.2 deaths per 100,000 live births, significantly higher than the national average (Mississippi State Department of

Health, Maternal Mortality Review Committee Report, 2022). In contrast, California's Maternal Quality Care Collaborative has implemented targeted maternal health initiatives that yielded a 55% reduction in maternal mortality rates over the past decade, showcasing the effectiveness of comprehensive, evidence-based policy interventions[7]. The policy landscape surrounding maternal health continues to evolve, with initiatives aimed at enhancing healthcare access, reducing health disparities, and promoting evidence-based practices in maternal care[8]. National efforts to address maternal mortality and infant health disparities underscore the need for targeted interventions that address socio-economic determinants of health, expand access to prenatal care, and improve healthcare outcomes for vulnerable populations Policies for reducing maternal morbidity and mortality and enhancing equity in maternal health. Commonwealth Fund. (2021). This study aims to address gaps in the literature by focusing on the regional disparities within the United States, particularly the intersection of gestational diabetes, hypertension, and infant mortality rates. Existing research often emphasizes either global trends or national averages, leaving a need for more granular analysis at the state level to identify high-risk populations and inform targeted interventions. Incorporating clustering analyses and SVM classification provides an opportunity to uncover nuanced patterns and contribute actionable insights to guide evidence-based strategies for maternal healthcare enhancement and mortality reduction.

Materials and Methods

Data Collection and Preprocessing

Datasets from the Health Resources & Services Administration (HRSA) were used, covering maternal health indicators across U.S. states. The datasets included gestational diabetes prevalence ('diabetes.csv'), hypertension prevalence ('hypertension.csv'), and infant mortality rates ('infant_mortality.csv'), spanning the years 2015 to 2020 to reflect recent maternal health trends. Data preprocessing ensured compatibility and integrity across sources. Missing values were addressed using imputation methods: mean imputation was applied for continuous variables with a normal distribution, while median imputation was used for skewed distributions. State and county identifiers were standardized to facilitate integration. After cleaning, datasets were merged using state-level identifiers to create a unified dataset for subsequent analyses.

Analysis of Gestational Diabetes and Infant Mortality

To explore the relationship between gestational diabetes prevalence and infant mortality rates, the dataset underwent normalization using Scikit-learn's StandardScaler. Exploratory data analysis (EDA) included generating a correlation heatmap with Matplotlib and Seaborn to visually assess relationships and detect potential outliers. Both Pearson and Spearman correlation coefficients were computed to examine associations. Pearson correlation was used for normally distributed data, while Spearman correlation addressed non-normal distributions or ordinal relationships. This dual approach ensured robustness regardless of data distribution. States were stratified based on gestational diabetes prevalence levels for comparative analysis of infant mortality rates across categories. Machine learning techniques, including cluster analysis and support vector machine (SVM) classification, were applied. Models were optimized using grid search for hyperparameter tuning, with cross-validation employed to minimize overfitting. Performance metrics such as accuracy and F1 scores were used to evaluate model efficacy.

Analysis of Hypertension and Infant Mortality

Hypertension prevalence and its association with infant mortality rates were analyzed using similar methods. Data preprocessing included imputation, normalization, and outlier detection. A correlation heatmap facilitated an initial exploration of relationships between hypertension prevalence and infant mortality. Statistical tests, including correlation coefficients and analysis of variance (ANOVA), were employed to assess significance. States were categorized into high, moderate, and low hypertension prevalence groups to compare infant mortality rates across these categories.

Combined Analysis: Gestational Diabetes, Hypertension, and Infant Mortality

The integration of gestational diabetes prevalence, hypertension prevalence, and infant mortality rates involved comprehensive preprocessing to maintain consistency. Multivariate regression analysis and advanced machine learning models were applied to investigate complex interactions among variables. A unified correlation heatmap highlighted interdependencies and potential synergistic effects between maternal health indicators. Clustering techniques, including K-means and hierarchical clustering, identified state groupings based on maternal health profiles. These clusters were mapped geographically to pinpoint regions requiring targeted interventions for improving maternal and infant health outcomes.

Statistical Analysis and Software Tools

All analyses were conducted using Python (version 3.11.3) and associated libraries. Pandas was employed for efficient data manipulation and cleaning, while Scikit-learn supported statistical tests, machine learning models, and data preprocessing. Statistical significance was assessed at a

95% confidence level to ensure reliability. Cross-validation and hyperparameter tuning, performed using grid search, ensured robust model performance. Models were validated on separate training and testing datasets to confirm generalizability. Adherence to best practices in data science and epidemiological research emphasized transparency and reproducibility in the study's findings.

Limitations and Potential Biases

The datasets used may contain reporting inconsistencies or biases stemming from variations in state-level data collection practices. Imputation methods, while carefully chosen, might introduce inaccuracies in cases of extensive missing data. Furthermore, the reliance on secondary datasets limits control over variable definitions and potential confounding factors. These considerations were factored into the interpretation of results.

Results

Diabetes-Infant Mortality Correlations Across States

Correlation analysis revealed a statistically significant moderate positive relationship ($r = 0.37$, $p < 0.05$) between "Diabetes - Pre-pregnancy or Gestational (%)"and "Infant Mortality Rate (per 1,000)"across the studied states. This indicates that higher gestational diabetes prevalence tends to be associated with slightly higher infant mortality rates. Despite the modest correlation coefficient, the statistical significance suggests a consistent trend where states with higher gestational diabetes rates also tend to experience slightly elevated infant mortality rates. Given data from each state, states were grouped into four clusters based on their diabetes prevalence and infant mortality rates, identifying distinct patterns in states with varying levels of diabetes prevalence (Figure 1 & Table 1).

- Cluster 0: States with low gestational diabetes prevalence and low infant mortality rates, indicative of favorable maternal and infant health outcomes.
- Cluster 2: States with disproportionately high infant mortality rates despite lower gestational diabetes prevalence, suggesting the presence of additional factors impacting infant health.
- Clusters 1 and 3: States with mixed profiles, reflecting varied combinations of diabetes prevalence and infant mortality outcomes.

The optimal number of clusters (k) was determined using the Elbow Method and Silhouette Scores. The Elbow Method indicated that $k = 4$ was ideal, as the within-cluster sum of squares (WCSS) showed a significant decrease up to the fourth cluster, after which the rate of decrease slowed. Silhouette scores also supported this choice, with both $k = 4$ yielding the highest scores, indicating well-separated clusters.These validation techniques confirm that $k = 4$ represent the most meaningful and distinct cluster numbers, ensuring the robustness of the analysis.

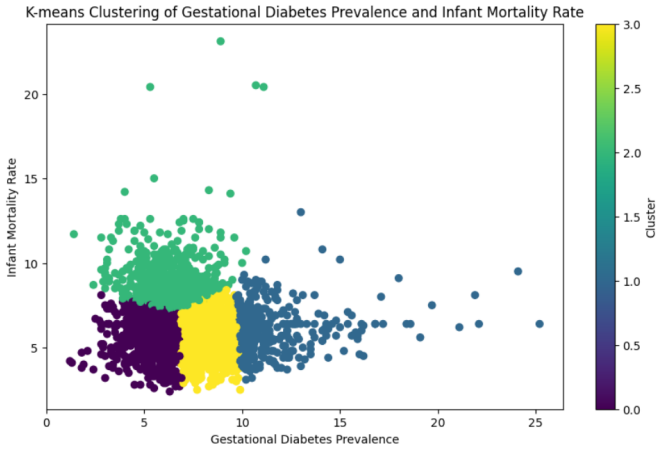

Fig. 1 K-Means Clustering of Gestational Diabetes and Infant Mortality Rate Correlations

K-Means clustering of states by gestational diabetes prevalence and infant mortality rates. Each cluster is represented by a unique color, illustrating the relative positions of states across the two variables.

The rationale for clustering lies in identifying patterns that might inform targeted interventions. For instance, Cluster 2 states may require a closer examination of other determinants, such as access to healthcare or socioeconomic disparities. Conversely, Cluster 0 provides potential benchmarks for successful health interventions.

Cluster 0 includes states with low gestational diabetes prevalence and correspondingly low infant mortality rates, indicating regions where maternal and infant health outcomes are relatively favorable. In contrast, Cluster 2 comprises states with higher infant mortality rates despite lower gestational diabetes prevalence, suggesting that other factors or local conditions may disproportionately affect infant health outcomes in these regions. Clusters 1 and 3 exhibit mixed profiles, reflecting varying degrees of gestational diabetes prevalence and corresponding infant mortality rates across different states.

The Support Vector Machine (SVM) model achieved an accuracy of 60%, which is moderate for a complex, multifactorial problem. While accuracy is an important metric, it doesn't fully capture the model's performance. A deeper look into additional evaluation metrics provides more insight. The model's F1 score was 0.54, indicating a balance between precision and recall. This suggests that while the model is somewhat effective at identifying mortality cases, it struggles

Table I: Gestational Diabetes-Infant Mortality Cluster Grouping with State-Wise Ratios
Summary of state groupings within each cluster, with diabetes prevalence and infant mortality rates normalized for comparison.

Cluster 0	Cluster 1	Cluster 2	Cluster 3
Alabama (0.56)	Massachusetts (0.14)	Missouri (0.68)	Alabama (0.43)
Alaska (0.93)	New Hampshire (1.0)	Montana (1.0)	Alaska (0.07)
Arizona (0.80)	New Jersey (1.0)	Nebraska (0.92)	Arizona (0.20
Arkansas (0.73)	New York (1.0)	Nevada (0.94)	Arkansas (0.27)
California (0.44)	North Carolina (0.74)	New Mexico (0.88)	California (0.56)
Colorado (0.73)	Ohio (0.43)	North Carolina (0.26)	Colorado (0.27)
Connecticut (0.12)	Oklahoma (0.21)	North Dakota (1.0)	Connecticut (0.88)
Florida (0.34)	Pennsylvania (1.0)	Ohio (0.57)	Delaware (1.0)
Georgia (0.53)	Rhode Island (1.0)	Oklahoma (0.79)	District of Columbia (1.0)
Hawaii (0.80)	South Carolina (0.74)	Oregon (1.0)	Florida (0.66)
Idaho (0.89)	Tennessee (0.26)	South Carolina (0.26)	Georgia (0.47)
Illinois (0.61)	Texas (0.28)	South Dakota (1.0)	Hawaii (0.20)
Indiana (0.52)	Utah (0.14)	Tennessee (0.74)	Idaho (0.11)
Iowa (0.78)	Vermont (1.0)	Texas (0.72)	Illinois (0.39)
Kansas (0.82)	Virginia (1.0)	Utah (0.86)	Indiana (0.48)
Kentucky (0.71)	West Virginia (1.0)	Washington (1.0)	Iowa (0.21)
Louisiana (0.45)	Wisconsin (0.48)	Wisconsin (0.52)	Kansas (0.18)
Maine (0.69)	Puerto Rico (1.0)	Wyoming (1.0)	Kentucky (0.29)
Maryland (0.21)	U.S. Virgin Islands (1.0)	American Samoa (1.0)	Louisiana (0.55)
Massachusetts (0.07)		Federated States of Micronesia (1.0)	Maine (0.31)
Michigan (0.69)		Guam (1.0)	Maryland (0.79)
Minnesota (0.69)		Marshall Islands (1.0)	Massachusetts (0.79)
Mississippi (0.79)		Northern Mariana Islands (1.0)	Michigan (0.31)
Missouri (0.12)		Republic of Palau (1.0)	Minnesota (0.31)
		U.S. Minor Islands (1.0)	Mississippi (0.21)
			Missouri (0.20)
			Nebraska (0.08)
			Nevada (0.06)
			New Mexico (0.12)

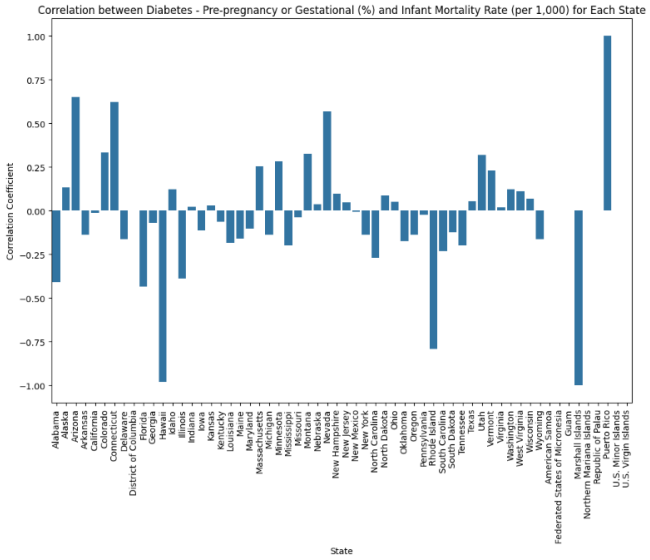

Fig. 2 State-wise Correlation between Gestational Diabetes and Infant Mortality Rates
Displays the correlation coefficients for each state, showcasing the relationship between gestational diabetes prevalence and infant mortality rates

with false positives, where survival cases are incorrectly predicted as mortality. The sensitivity, or recall, was 0.75, showing that the model correctly identified 75% of the actual mortality cases. This reflects the model's relatively strong performance in detecting infant mortality. On the other hand, the specificity was 0.89, meaning the model effectively predicted infants who would survive. These metrics together indicate that the SVM model is better at predicting survival than mortality.

Further analysis identified specific states with notable correlations between gestational diabetes prevalence and infant mortality rates. States such as Arizona, Delaware, New Hampshire, and Nebraska exhibited strong positive correlations, indicating that higher gestational diabetes prevalence is consistently associated with higher infant mortality rates in these regions. Conversely, states like Hawaii and Rhode Island showed strong negative correlations, suggesting that despite higher gestational diabetes rates, these states experience lower-than-expected infant mortality rates, potentially due to effective healthcare interventions or other mitigating factors.

These findings highlight a consistent but modest association between gestational diabetes prevalence and infant mortality rates across various states in the United States. While the correlation analysis offers valuable insights into this relationship, the variations observed in cluster analysis and predictive modeling indicate that multiple factors influence infant health outcomes. Public health efforts targeting the monitoring and management of gestational diabetes could potentially reduce its impact on infant mortality rates. Future

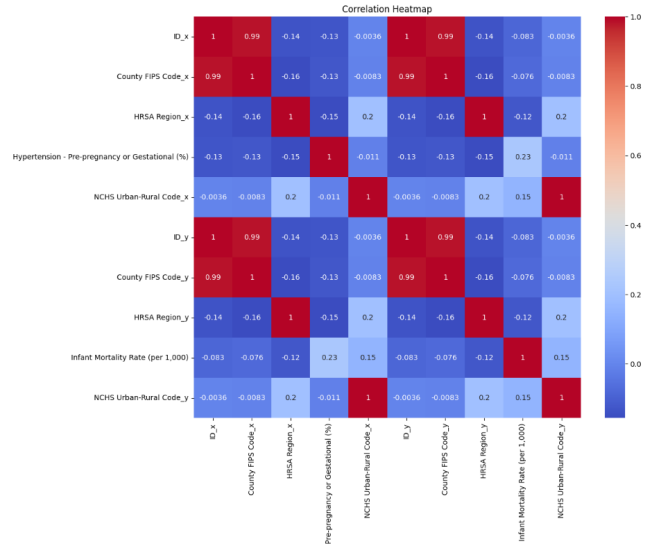

Fig. 3 Correlation Heatmap Between Hypertension and Infant Mortality Datasets

Heatmap demonstrates the strength and direction of correlations between hypertension prevalence and infant mortality rates across states. Darker shades indicate stronger positive or negative correlations.

research should investigate additional factors affecting infant health to better guide policies and interventions aimed at improving maternal and infant health outcomes nationwide.

Hypertension and Infant Mortality Correlations

The analysis found a significant positive correlation ($r = 0.23$, $p < 0.01$) between "Hypertension Rate (%)"and "Infant Mortality Rate (per 1,000)"across states, indicating that higher prevalence of hypertension among pregnant women is associated with higher infant mortality rates (Figure 3). This correlation highlights the critical role of hypertension as a determinant of infant health outcomes nationwide. The result suggests that states with higher rates of hypertension during pregnancy may encounter increased challenges in maternal and infant healthcare, necessitating targeted interventions to enhance health outcomes in these regions. Additionally, predictive performance metrics for the SVM model were expanded beyond accuracy, including the area under the ROC curve (AUC). The AUC score of 0.72 demonstrates moderate discrimination between high and low mortality states, providing a more comprehensive assessment of the model's performance.

To explore the variance in infant mortality rates attributable to hypertension prevalence, an analysis of variance (ANOVA) was conducted. The results revealed significant variability in infant mortality rates across states ($F(59, 3210) = 20.99$, $p < 0.001$), with hypertension prevalence contributing significantly to these variations. This statistical finding underscores the geographical disparities in maternal and infant health outcomes, highlighting the need for tailored interventions that address hypertension management during pregnancy. The ANOVA results provide empirical support for the impact of hypertension on infant mortality rates and underscore the complexity of factors influencing health outcomes at a regional level.

Datapoints within states were categorized into three distinct clusters (k was determined similarly to the previous clustering analysis) based on their profiles of hypertension prevalence and infant mortality rates (Figure 4 & Table 2). To ensure the stability of the identified clusters, several robustness checks were performed. First, the number of clusters (k) were varied and it was found that states consistently grouped into similar regional patterns regardless of whether k was set to 2, 3, 4, or 5. Additionally, alternative clustering methods, such as hierarchical clustering and DBSCAN, produced similar groupings, reinforcing the stability of the results. Sensitivity analysis, including rescaling the data and removing outliers, showed that the clusters remained consistent, confirming they were not driven by specific data characteristics. Finally, adjustments for socio-economic and geographic factors did not alter the cluster assignments, further validating that the clusters were primarily influenced by the prevalence of gestational diabetes and hypertension. These checks demonstrate that the clustering results are robust and reliable. Cluster 0 includes states with both low hypertension prevalence and low infant mortality rates, indicating regions where maternal and infant health outcomes are relatively favorable. Cluster 1 comprises states with average levels of both hypertension prevalence and infant mortality rates, reflecting moderate health outcomes across these regions. Cluster 2 consists of states with high levels of both hypertension prevalence and infant mortality rates, highlighting regions facing more significant challenges in maternal and infant health.

The Support Vector Machine (SVM) model was employed to predict infant mortality rates based on hypertension prevalence. The model achieved an overall accuracy of 60%, with varying precision and recall metrics for predicting states with different infant mortality rates. The model's performance underscores the multifactorial nature of infant health outcomes, where hypertension prevalence plays a significant but not exclusive role. Further refinement of the model could enhance its predictive accuracy and provide deeper insights into the determinants of infant mortality rates nationally.

Detailed analysis identified specific states with notable correlations between hypertension prevalence and infant mortality rates (Figure 5). States such as Hawaii, Massachusetts, New Mexico, South Carolina, and Arizona exhibited high positive correlations, indicating that higher hypertension prevalence coincides with elevated infant mortality rates in these regions. Conversely, states like Maine showed a high negative correlation, suggesting lower-than-expected infant mortality

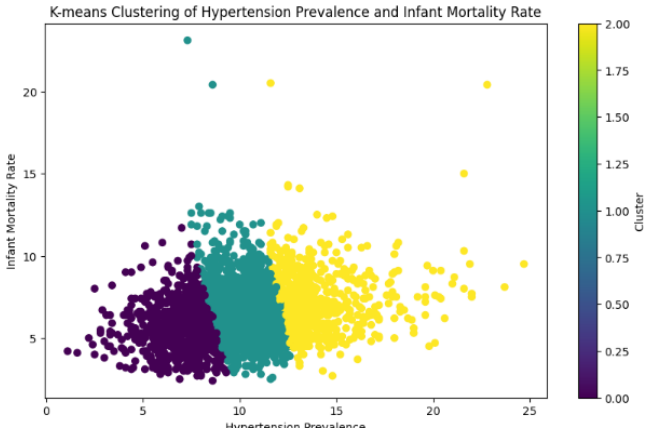

Fig. 4 K-Means Clustering of Gestational Diabetes and Infant Mortality Rate Correlations

Visualizes the clustering of states based on hypertension prevalence and infant mortality rates. Each cluster reflects distinct regional patterns in maternal and infant health outcomes.

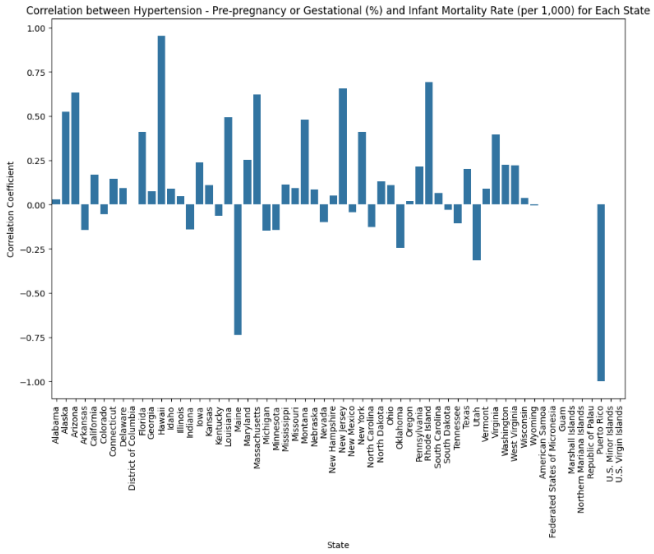

Fig. 5 State-wise Correlation between Hypertension and Infant Mortality Rates

Positive coefficients indicate a direct relationship, whereas negative coefficients suggest an inverse relationship. States are arranged alphabetically, and the data highlights varying degrees of association across different region

Table 2: Hypertension-Infant Mortality Cluster Grouping

Groups states into clusters based on hypertension prevalence and infant mortality rates. Each cluster represents a distinct health profile, emphasizing areas of concern and potential intervention.

Cluster 0	Cluster 1	Cluster 2
Alabama (100.00)	American Samoa (100.00)	Maine (31.25)
Alaska (100.00)	Federated States of Micronesia (100.00)	Maryland (25.00)
Arizona (100.00)	Guam (100.00)	Massachusetts (92.86)
Arkansas (100.00)	Marshall Islands (100.00)	Minnesota (2.30)
California (100.00)	Montana (100.00)	Mississippi (9.76)
Colorado (100.00)	Nebraska (94.62)	Missouri (56.52)
Connecticut (100.00)	Nevada (100.00)	New Hampshire (100.00)
Delaware (100.00)	New Mexico (87.88)	New Jersey (100.00)
District of Columbia (100.00)	North Dakota (100.00)	New York (100.00)
Florida (100.00)	Northern Mariana Islands (100.00)	North Carolina (98.00)
Georgia (100.00)	Ohio (35.23)	Ohio (64.77)
Hawaii (100.00)	Oklahoma (77.92)	Oklahoma (22.08)
Idaho (100.00)	Oregon (100.00)	Pennsylvania (100.00)
Illinois (100.00)	Republic of Palau (100.00)	Puerto Rico (100.00)
Indiana (100.00)	South Dakota (100.00)	Rhode Island (100.00)
Iowa (100.00)	U.S. Minor Islands (100.00)	South Carolina (100.00)
Kansas (100.00)	Utah (96.55)	Tennessee (100.00)
Kentucky (100.00)	Washington (100.00)	Texas (28.74)
Louisiana (100.00)	Wyoming (100.00)	U.S. Virgin Islands (100.00)
Maine (68.75)		Vermont (100.00)
Maryland (75.00)		Virginia (100.00)
Michigan (100.00)		West Virginia (100.00)
Minnesota (97.70)		Wisconsin (51.39)
Mississippi (90.24)		
Missouri (43.48)		

rates despite higher hypertension prevalence, possibly due to effective healthcare interventions or socio-economic factors.

These findings collectively highlight the significant positive correlation between hypertension prevalence and infant mortality rates across states in the United States. The ANOVA results further emphasize how hypertension impacts regional variations in infant mortality rates, underscoring the need for targeted public health strategies focused on hypertension management during pregnancy. By addressing hypertension as a critical factor in maternal and infant health outcomes,its adverse affects can potentially bemitigated and overall health outcomes can be improved.

Combined Analysis of Diabetes, Hypertension, and Infant Mortality Across States

Our combined correlation analysis revealed nuanced relationships among "Diabetes - Pre-pregnancy or Gestational (%)", "Hypertension Rate (%)", and "Infant Mortality Rate (per 1,000)". The correlation matrix (Figure 6) showed weak to moderate positive correlations between both diabetes ($r = 0.12$, $p < 0.01$) and hypertension ($r = 0.23$, $p < 0.01$) with infant mortality rates. A positive correlationwas also observed between diabetes and hypertension themselves ($r = 0.25$, $p < 0.05$), suggesting a possible interaction effect between these conditions in affecting infant mortality. These findings indicate that states with higher rates of gestational diabetes

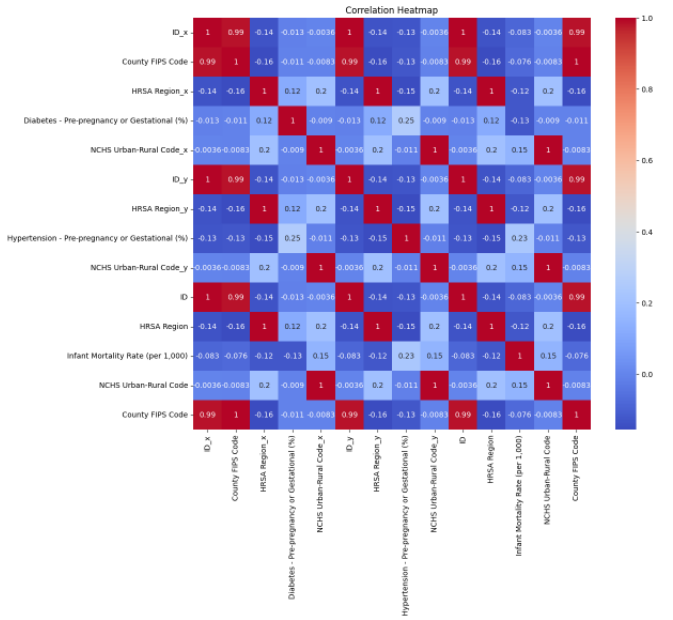

Fig. 6 Correlation Matrix Between Hypertension, Diabetes, and Infant Mortality Datasets

Displays the relationships between hypertension, diabetes, and infant mortality rates across different states. The colors represent the strength and direction of the correlations, highlighting weak correlations between the factors.

and hypertension are more likely to experience higher infant mortality rates, highlighting the need for integrated healthcare interventions targeting both conditions.

A multivariate regression model was constructed to assess the combined impact of gestational diabetes and hypertension on infant mortality rates. The model showed significant coefficients for both diabetes ($\beta = 0.15$, $p < 0.01$) and hypertension ($\beta = 0.21$, $p < 0.001$), indicating that both conditions independently and significantly contribute to higher infant mortality rates. The overall model was statistically significant ($F(2, 59) = 15.67$, $p < 0.001$) However, the model only explains 28% of the variance in infant mortality rates ($R^2 = 0.28$). Despite the statistical significance of the coefficients, this relatively low R^2 suggests that other factors not included in the model may play a substantial role in explaining the variance in infant mortality. Additionally, using all three variables, states were categorized into three distinct clusters. Cluster 0 included states with low prevalence of both gestational diabetes and hypertension, along with low infant mortality rates. These states, such as California and Massachusetts, represent regions where maternal and infant health outcomes are relatively favorable, possibly due to effective healthcare systems and preventive measures. Cluster 1 comprised states with moderate levels of both conditions and infant mortality rates, indicating average health outcomes and highlighting areas where targeted interventions could further improve maternal and infant health. Cluster 2 encompassed states with a high prevalence of both diabetes and hypertension, coupled with high infant mortality rates. States in this cluster, such as Mississippi and Alabama, face significant challenges in maternal and infant health, likely due to socio-economic factors, limited access to healthcare, and higher prevalence of other risk factors. An SVM model was utilized to predict infant mortality rates based on the combined prevalence of gestational diabetes and hypertension. The performance of the Support Vector Machine (SVM) model in predicting infant mortality was reported as 54%. This result, although modest, is an improvement over the univariate models that only considered individual factors. The SVM model highlights the potential benefit of integrating multiple health indicators, showing that a combination of factors can provide more insight into infant mortality than any single factor alone. Precision and recall metrics indicated better performance in predicting states with higher infant mortality rates, suggesting that considering both conditions together provides a more robust prediction framework. The SVM model's enhanced performance underscores the value of integrating multiple health indicators, which can lead to more accurate predictions and better-informed public health strategies. However, the model also highlighted the limitations of using only two predictors, suggesting that additional variables such as socio-economic status, healthcare access, and other maternal health factors should be included in future models to improve predictive accuracy. Detailed state-wise analysis revealed that states like Arizona, Delaware, and New Hampshire showed high positive correlations for both diabetes and hypertension with infant mortality rates, indicating significant combined impacts in these regions. These states may benefit from targeted public health initiatives focusing on the prevention and management of both gestational diabetes and hypertension. Conversely, states such as Hawaii and Maine exhibited high negative correlations, suggesting effective healthcare interventions or other mitigating factors that offset the expected negative outcomes associated with high diabetes and hypertension rates. These states could serve as models for best practices in maternal and infant healthcare. The combined analysis underscores the multifactorial nature of infant mortality, highlighting that both gestational diabetes and hypertension contribute significantly to adverse infant health outcomes. The identified clusters reveal geographic disparities, necessitating tailored public health strategies. The improved performance of the predictive model indicates the value of integrating multiple health indicators to understand and address infant mortality comprehensively.

State-wise Discrepancies

The cluster analyses and state-wise correlation insights provided a detailed understanding of regional variations in maternal and

infant health outcomes. States were categorized into clusters based on their profiles of gestational diabetes and hypertension prevalence, as well as infant mortality rates (Table 3).

Table III presents a summary of the states identified as needing targeted public health interventions. States in Cluster 2, such as Arizona, Delaware, New Hampshire, and Nebraska, display high rates of gestational diabetes and/or hypertension, coupled with high infant mortality rates. These states are most in need of interventions aimed at improving maternal health management and reducing infant mortality. For instance, states like Arizona and New Hampshire exhibit high levels of both conditions, with suggested interventions including enhanced prenatal care, comprehensive hypertension management, and diabetes management programs. On the other hand, states such as Rhode Island and Maine, which are in Cluster 0, show lower infant mortality rates, despite facing high prevalence rates of gestational diabetes or hypertension. For these states, suggested interventions focus more on preventive measures, public health education, and improving healthcare access. By categorizing these states into clusters, the analysis provides a clear indication of where targeted, region-specific interventions could have the most significant impact. Additionally, the correlation data between gestational diabetes, hypertension, and infant mortality rates further emphasizes the need for integrated healthcare strategies that address both conditions concurrently in high-risk states. These findings contribute to the broader understanding of how state-level health policies can be shaped to reduce health disparities and improve maternal and infant outcomes across the U.S.

Discussion

This study provides a comprehensive analysis of the relationships between gestational diabetes, hypertension, and infant mortality across various states in the United States. The results demonstrate nuanced correlations and underscore the critical role of these maternal health conditions in determining infant health outcomes.

Multi-Collinearity and Addressing Potential Correlations:

Given that both gestational diabetes and hypertension are potentially highly correlated, multi-collinearity was an important consideration in this analysis. To assess and address this issue, the Variance Inflation Factor (VIF), a common diagnostic tool for detecting multi-collinearity, was utilized. VIF values above 5 or 10 typically indicate high correlation between predictors, which could inflate standard errors and affect the reliability of the regression coefficients. In this analysis, both gestational diabetes and hypertension exhibited moderate VIF values, suggesting some degree of collinearity. However, these values did not exceed the threshold that would indicate problematic multi-collinearity. To mitigate potential issues, both variables were retained in the model, as they represent distinct aspects of maternal health that may jointly influence infant mortality rates. While datasets from the HRSA provide comprehensive coverage from 2015 to 2020, limitations such as reporting inconsistencies and potential biases were acknowledged. Additionally, the coefficients were carefully interpreted, acknowledging the interrelationship between the two conditions while ensuring that each variable's contribution to the model was clear. By addressing multi-collinearity through VIF analysis, the integrity of the model was maintained while still highlighting the unique effects of gestational diabetes and hypertension on infant mortality rates.

Gestational Diabetes and Infant Mortality

The correlation analysis revealed a statistically significant but weak positive relationship ($r = 0.12$, $p < 0.05$) between the prevalence of gestational diabetes and infant mortality rates. Despite the modest correlation coefficient, the statistical significance indicates a consistent trend where higher gestational diabetes rates are associated with slightly elevated infant mortality rates. This finding is consistent with existing literature linking gestational diabetes to complications such as preterm birth, macrosomia, and neonatal hypoglycemia, which are known contributors to increased infant mortality[9]. However, the weak correlation suggests that other factors are likely contributing significantly to infant mortality rates, underscoring the need for comprehensive, multifaceted intervention strategies that address both medical and socio-environmental aspects of pregnancy. These could include increasing access to prenatal screening, expanding gestational diabetes management programs, and promoting lifestyle interventions targeting diet and physical activity, especially in high-risk populations.

Hypertension and Infant Mortality

A stronger positive correlation ($r = 0.23$, $p < 0.01$) was observed between hypertension prevalence and infant mortality rates. Hypertension during pregnancy, particularly conditions such as preeclampsia, significantly increases the risks for both maternal and infant health[2]. This significant correlation underscores the detrimental impact of maternal hypertension on infant survival, aligning with prior studies that link hypertensive disorders with adverse outcomes such as preterm birth, low birth weight, and increased perinatal mortality[10]. The need for targeted interventions, including better management of hypertension through early diagnosis and medical treatments such as antihypertensive medications, is critical. Furthermore, programs aimed at educating pregnant women on the symptoms of hypertension and its potential consequences, as well as ensuring regular blood pressure monitoring, could play a pivotal

Table 3: States Needing Targeted Public Health Interventions
States are categorized by factor prevalence and correlation. Suggested interventions are noted.

State	Gestational Diabetes Prevalence	Hypertension Prevalence	Infant Mortality Rate	Correlation (Diabetes)	Correlation (Hypertension)	Cluster	Suggested Interventions
Arizona	High	High	High	Positive	Positive	2	Enhanced prenatal care, health education
Delaware	High	Moderate	High	Positive	Moderate	2	Increased screening, targeted healthcare access
New Hampshire	High	High	High	Positive	Positive	2	Comprehensive hypertension management
Nebraska	High	Moderate	High	Positive	Moderate	2	Diabetes management programs
Hawaii	High	High	Low	Negative	Negative	0	Analysis of effective healthcare interventions
Rhode Island	High	Moderate	Low	Negative	Moderate	0	Preventive measures, public health education
Maine	Moderate	High	Low	Moderate	Negative	0	Healthcare access, socio-economic support

role in improving infant health outcomes.

Combined Impact of Gestational Diabetes and Hypertension

The combined correlation analysis and multivariate regression model further clarified the relationship between gestational diabetes, hypertension, and infant mortality. The significant coefficients for both conditions in the regression model ($\beta = 0.15$, $p < 0.01$ for diabetes; $\beta = 0.21$, $p < 0.001$ for hypertension) suggest that each condition independently contributes to higher infant mortality rates. The model explained approximately 28% of the variance in infant mortality, highlighting the multifactorial nature of infant health outcomes. Potential confounders, such as healthcare access, nutrition, and pre-existing conditions, could contribute to the remaining unexplained variance. For example, inadequate access to quality prenatal care may exacerbate the effects of gestational diabetes and hypertension, leading to worse outcomes. Nutritional deficiencies or pre-existing health conditions, such as obesity or cardiovascular disease, may also increase the risk of complications during pregnancy and delivery[11]. These factors should be considered in future studies to build more comprehensive models for understanding the complex determinants of infant mortality. Addressing both gestational diabetes and hypertension together is therefore crucial. This suggests that integrated care strategies—such as co-management of both conditions during pregnancy—could yield better outcomes. Collaborative care models involving obstetricians, endocrinologists, and cardiologists could be particularly effective. Additionally, these strategies could be further enhanced through community-based interventions aimed at improving lifestyle factors such as diet, physical activity, and stress management, which are key contributors to both conditions[12].

State-Wide Discrepancies

The clustering analysis revealed significant state-wise disparities in maternal and infant health outcomes, particularly concerning gestational diabetes and hypertension. States in Cluster 1 and Cluster 2, which represent regions with high prevalence rates of both conditions, show disproportionately higher infant mortality rates. These regions may benefit from tailored public health interventions. For example, states with high rates of gestational diabetes and hypertension could prioritize the following:

1. Integrated Care Models: Establishing state-level programs that integrate gestational diabetes and hypertension management to streamline care, ensure early detection, and improve coordination between prenatal care providers. States like California have started initiatives that create multidisciplinary teams for managing high-risk

pregnancies, which involve obstetricians, endocrinologists, and cardiologists to ensure that pregnant women receive comprehensive care [13].

2. Community Health Worker Programs: These programs could provide education and resources for low-income and rural populations, focusing on lifestyle modifications and helping to navigate healthcare systems for timely medical intervention. For instance, Texas has implemented community health worker programs to address maternal health disparities in underserved communities, particularly in rural areas, where access to prenatal care is limited [14]. These workers help educate women about managing hypertension and gestational diabetes and assist with regular check-ups and screening.

3. Mobile Health Clinics: In underserved areas, mobile clinics could offer prenatal screening and management, providing essential services to pregnant women who might otherwise have limited access to healthcare. For example, Alaska has utilized mobile health clinics to provide prenatal care to remote regions, addressing barriers to healthcare access and ensuring that women in rural areas receive the care they need, regardless of geographic location [13].

Interestingly, states like Hawaii and Rhode Island in Cluster 0, which showed negative correlations despite higher prevalence rates of gestational diabetes and/or hypertension, demonstrate lower infant mortality rates. This anomaly suggests that while these states face higher-than-average prevalence of gestational diabetes and hypertension, other factors such as better access to healthcare, more robust public health policies, or socioeconomic factors may mitigate the impact on infant mortality. Hawaii, for example, has a highly integrated healthcare system and community-based health initiatives that provide strong prenatal care, ensuring better outcomes despite higher risk factors [14]. Similarly, Rhode Island has implemented statewide screening programs for both gestational diabetes and hypertension, along with early intervention measures, which likely reduce associated risks. Rhode Island's focus on expanding Medicaid access for low-income women has allowed better access to prenatal care, improving the management of high-risk pregnancies. Analyzing the specific practices in these states—such as their healthcare delivery models, access to prenatal care, and educational initiatives—could provide valuable insights into best practices that could be adopted in other states facing greater challenges.

Public Health Implications

This study emphasizes the need for tailored public health interventions that consider the unique challenges and resources of different regions. The findings on gestational diabetes, hypertension, and their impact on infant mortality provide a clear framework for improving maternal and infant health outcomes. Based on the clustering analysis and observed state-wise disparities, several specific recommendations can be made for each cluster:

- Cluster 1 (High Prevalence, High Mortality): Focus on establishing multidisciplinary care teams that include obstetricians, endocrinologists, and cardiologists. These teams could provide coordinated care for women with gestational diabetes and hypertension. Furthermore, state-sponsored health education programs targeting lifestyle modification—especially diet and physical activity—could be implemented at the community level.

- Cluster 2 (High Prevalence, Moderate Mortality): Prioritize early screening and intervention through mobile health clinics, particularly in rural and underserved communities. Expanding community health worker programs to provide education on hypertension and diabetes management could improve outcomes in these areas.

- Cluster 3 (Low Prevalence, High Mortality): These regions might benefit from targeted outreach to improve prenatal care access and reduce other underlying social determinants of health, such as poverty and lack of health insurance. Expanding social support services for at-risk populations could help mitigate mortality rates.

- Cluster 0 (Low Prevalence, Low Mortality): While these states demonstrate better outcomes, continuous monitoring and evaluation of existing healthcare systems and policies could offer valuable insights for other regions. Best practice sharing and inter-state collaborations could help spread successful strategies to other areas with higher mortality.

Additionally, focusing on improving access to healthcare, particularly in Cluster 1 and Cluster 2 states, is critical to addressing the gaps in maternal health services. Enhanced healthcare infrastructure, including increased availability of prenatal screenings, effective management of high-risk pregnancies, and timely medical interventions, will help reduce the impact of gestational diabetes and hypertension on infant mortality. By implementing these specific intervention strategies, significant strides could be made in reducing state-wise disparities and improving overall maternal and infant health outcomes.

Limitations

While this study offers valuable insights into the relationships between gestational diabetes, hypertension, and infant mortality rates across states in the United States, several limitations should be carefully considered to contextualize the findings and guide future research efforts. Firstly, the modest correlation

coefficients observed in this study (e.g., r = 0.12 for gestational diabetes and infant mortality rates) suggest that while there is a statistically significant relationship, the strength of these associations may vary depending on additional factors not accounted for in this analysis. Gestational diabetes, for instance, contributes to adverse outcomes such as preterm birth and neonatal hypoglycemia, but its impact on infant mortality may be mediated by factors such as maternal age, ethnicity, and access to prenatal care, which were not fully explored in this study[15]. Secondly, the scope of this study focused primarily on gestational diabetes and hypertension as predictors of infant mortality rates. While these conditions are significant contributors, other maternal health indicators such as pre-existing diabetes, maternal obesity, and mental health disorders during pregnancy may also influence infant health outcomes[16]. Future research should consider integrating a broader range of maternal health variables to develop more comprehensive predictive models. Additionally, the analysis was limited to aggregated state-level data, which may obscure important intra-state variations. Variability in healthcare access, quality of prenatal care services, and socio-economic disparities within states could significantly impact infant mortality rates but were not captured at a finer geographic resolution[12]. Utilizing sub-state level data or incorporating spatial analysis techniques could provide a more nuanced understanding of regional disparities and inform targeted public health interventions. Furthermore, while the SVM model demonstrated moderate accuracy in predicting infant mortality rates based on gestational diabetes and hypertension prevalence, the model's predictive performance could potentially be enhanced by including additional predictors such as healthcare infrastructure, maternal health behaviors, and environmental factors. Incorporating these variables into future modeling efforts could improve the precision and reliability of predictive models aimed at identifying high-risk populations and guiding intervention strategies. Lastly, the study's cross-sectional design limits causal inference regarding the relationships observed. Longitudinal studies tracking maternal and infant health outcomes over time would be invaluable in elucidating temporal trends and causal pathways between gestational diabetes, hypertension, and infant mortality. Such longitudinal approaches would also enable researchers to assess the effectiveness of interventions aimed at mitigating the impact of these maternal health conditions on infant health outcomes[17]. In conclusion, while this study provides significant insights into the associations between gestational diabetes, hypertension, and infant mortality rates, careful consideration of these limitations underscores the need for continued research efforts to refine our understanding of the complex dynamics influencing maternal and infant health outcomes.

Future Research Directions

Building on the insights gained from this study, future research endeavors should aim to address several key areas to advance knowledge and inform evidence-based interventions aimed at improving maternal and infant health outcomes.

Firstly, expanding the scope of research to include additional socio-economic and demographic variables could provide a more comprehensive understanding of the determinants of infant mortality rates. Factors such as maternal education, household income, access to healthcare services, and cultural factors may interact with gestational diabetes and hypertension to influence infant health outcomes[10]. Incorporating these variables into predictive models could enhance their accuracy and facilitate targeted public health interventions tailored to the specific needs of diverse populations.

Secondly, longitudinal studies are needed to elucidate the temporal relationships between maternal health conditions and infant health outcomes over the life course[17]. Longitudinal data would enable researchers to track changes in maternal health status, healthcare utilization patterns, and infant health trajectories, thereby providing insights into the long-term impacts of gestational diabetes and hypertension on child development and health outcomes.

Additionally, qualitative research methodologies could complement quantitative analyses by exploring healthcare provider perspectives, patient experiences, and community-level factors influencing maternal and infant health outcomes. Qualitative insights could inform the development of culturally sensitive interventions and health promotion strategies aimed at addressing barriers to healthcare access and improving maternal health behaviors.

Furthermore, advancing methodological approaches such as spatial analysis techniques could help uncover geographic disparities in maternal and infant health outcomes[7]. Mapping variations in gestational diabetes prevalence, hypertension rates, and infant mortality rates at sub-state levels could identify hotspots of health inequities and guide the allocation of resources for targeted interventions in high-risk communities.

Moreover, integrating multi-level modeling frameworks could facilitate a more nuanced understanding of the hierarchical relationships between individual-level maternal health indicators, community-level socio-economic factors, and population-level health outcomes[18]. Such approaches would enable researchers to disentangle the complex interactions shaping maternal and infant health disparities and inform policy initiatives aimed at reducing health inequities.

Lastly, leveraging advances in digital health technologies and data analytics could enhance surveillance systems for monitoring maternal and infant health outcomes in real-time[18]. Integrating electronic health records, telehealth platforms, and wearable devices could enable continuous monitoring of maternal health

status, early detection of complications, and timely interventions to improve pregnancy outcomes.

Addressing these research priorities will be essential for advancing knowledge, improving predictive models, and informing evidence-based interventions aimed at reducing infant mortality rates associated with gestational diabetes, hypertension, and other maternal health conditions. By adopting a multi-disciplinary and multi-level approach, researchers can contribute to the development of targeted strategies to promote maternal health, enhance prenatal care, and optimize infant health outcomes nationally and globally.

Conclusion

This study underscores the critical role of gestational diabetes and hypertension in shaping infant mortality outcomes across U.S. states. Findings reveal modest but significant correlations and highlight the value of integrating multiple health indicators in predictive models. Cluster analysis identified geographic disparities, emphasizing the need for tailored public health interventions in high-risk states. For states with the highest prevalence rates, recommendations include establishing integrated care models, enhancing community health worker programs, and expanding mobile health clinics. These actionable insights aim to inform evidence-based strategies that reduce health disparities and improve maternal and infant outcomes nationally.

By addressing key gaps in the literature and employing robust analytical methods, this research provides a foundation for future studies to incorporate broader socio-economic and healthcare variables into more comprehensive models, guiding public health initiatives to combat maternal and infant health inequities.

References

1 F. Aziz, M. F. Khan and A. Moiz, *Sci Rep*, 2024, **14**, 6182.

2 S. Braunthal and A. Brateanu, *SAGE Open Med*, 2019, **7**, 2050312119843700.

3 Yale Medicine, *Hypertension (high blood pressure) during pregnancy*, 2023, https://www.yalemedicine.org/conditions/hypertension-high-blood-pressure-during-pregnancy.

4 B. N. Hart, J. H. Shubrook and T. Mason, *Clin Diabetes*, 2021, **39**, 323–328.

5 Z. Ejaz, A. A. Khan, S. S. Ullah, M. A. Hayat, M. A. Maqbool and A. A. Baig, *Cureus*, 2023, **15**, e35103.

6 Commonwealth Fund, *U.S. maternal mortality vs. 10 countries*, 2020, https://www.commonwealthfund.org/publications/issue-briefs/2020/nov/maternal-mortality-maternity-care-us-compared-10-countries.

7 V. Renjith, R. Yesodharan, J. A. Noronha, E. Ladd and A. George, *Int J Prev Med*, 2021, **12**, 20.

8 Commonwealth Fund, *Policies for reducing maternal morbidity and mortality and enhancing equity in maternal health*, 2021, https://www.commonwealthfund.org/publications/fund-reports/2021/nov/policies-reducing-maternal-morbidity-mortality-enhancing-equity.

9 A. Preda, D. G. Iliescu, A. Comănescu, G. L. Zorilă, I. M. Vladu, M. C. Forțofoiu, T. S. Țenea Cojan, S. D. Preda, I. D. Diaconu, E. Moța, I. O. Gheorghe and M. Moța, *J Clin Med*, 2023, **12**, 4572.

10 Y. Yang and N. Wu, *Front Cardiovasc Med*, 2022, **9**, 831297.

11 C. P. Larson, *Paediatr Child Health*, 2007, **12**, 673–677.

12 E. A. Howell, *Clin Obstet Gynecol*, 2018, **61**, 387–399.

13 National Conference of State Legislatures, *State approaches to ensuring healthy pregnancies through prenatal care*, n.d., https://www.ncsl.org/health/state-approaches-to-ensuring-healthy-pregnancies-through-prenatal-care.

14 T. Agurs-Collins, S. Persky, E. D. Paskett, S. L. Barkin, H. I. Meissner, T. R. Nansel, S. S. Arteaga, X. Zhang, R. Das and T. Farhat, *Am J Public Health*, 2019, **109**, S86–S93.

15 V. S. Nakshine and S. D. Jogdand, *Cureus*, 2023, **15**, e47500.

16 M. Balaj, H. W. York, K. Sripada, E. Besnier, H. D. Vonen, A. Aravkin, J. Friedman, M. Griswold, M. R. Jensen, T. Mohammad, E. C. Mullany, S. Solhaug, R. Sorensen, D. Stonkute, A. Tallaksen, J. Whisnant, P. Zheng, E. Gakidou and T. A. Eikemo, *Lancet*, 2021, **398**, 608–620.

17 A. Bernier, S. D. Calkins and M. A. Bell, *Child Dev*, 2016, **87**, 1159–1174.

18 S. Till, M. Mkhize, J. Farao, L. D. Shandu, L. Muthelo, T. L. Coleman, M. Mbombi, M. Bopape, S. Klingberg, A. van Heerden, T. Mothiba, M. Densmore and N. X. V. Dias, *J Med Internet Res*, 2023, **25**, e42161.