Sugar and Status

Glycaemic Risk and Social Stratification Among Women in India

NFHS-5 Data Analysis

India has a diabetes problem, and it is not the one that usually makes headlines. The conversation tends to focus on the country's rising middle class, on processed food consumption and sedentary urban lifestyles. These things are real. But they sit on top of a far more complex picture. The data from the National Family Health Survey's fifth round tell a story where blood glucose is not just a biological phenomenon. It is a social one.

This analysis examines plasma glucose levels, prediabetes and diabetes prevalence, and the associated social determinants among women aged 15 to 49 in India. The findings draw on NFHS-5 data using survey-weighted estimation and ordered logistic regression. What emerges is a pattern that resists simple interpretation and demands careful attention.

A Gradient That Defies Expectation

Start with wealth. Among the poorest women in the sample, mean plasma glucose sat at 120.06 mg/dL. Among the richest, it reached 140.39 mg/dL. That is a gap of over 20 mg/dL across wealth quintiles, and it runs in the direction that surprises most people who encounter it for the first time.

The poorest are supposed to carry the heaviest disease burden. In much of the world they do. But diabetes, at least in its early stages in India's demographic context, is concentrating at the top. The richest quintile had a diabetes prevalence of 6.2 per cent, compared to 2.6 per cent in the poorest group. This is not a fluke. It is consistent with what a sedentary lifestyle, high-calorie diet and reduced physical labour look like when they become concentrated by class.

Table 1: Mean Plasma Glucose and Diabetes Prevalence by Wealth Index (NFHS-5)

Wealth Quintile	Mean Glucose (mg/dL)	95% CI	Prediabetic (%)	Diabetic (%)
Poorest	120.06	119.14 - 120.98	7.4	2.6
Poorer	120.49	119.60 - 121.37	7.0	2.5
Middle	122.69	121.72 - 123.66	7.1	3.3
Richer	128.97	127.77 - 130.18	7.2	4.5
Richest	140.39	138.74 - 142.03	6.9	6.2

Source: NFHS-5; one-way ANOVA, $F(4, 210818) = 205.17$, $p < 0.001$

Notice what the prediabetes column does. It stays almost flat across the wealth quintile spectrum, hovering around 7 per cent throughout. Diabetes prevalence, by contrast, more than doubles between the poorest and richest groups. What this suggests is that prediabetes is broadly distributed while the progression to full diabetes is shaped by factors that cluster among the wealthy. Sedentary office jobs, highly caloric diets, stress without physical release. The conversion from at-risk to diabetic is not random.

The Urban Divide

Urban women in the sample had a mean plasma glucose of 137.40 mg/dL. Rural women sat at 122.04 mg/dL. That is a 15.4 mg/dL gap, which is one of the largest single differentials in the entire analysis ($t = 26.10$, $p < 0.001$). Urban women had nearly twice the diabetes prevalence of their rural counterparts, 5.6 per cent against 3.0 per cent.

But here is where the story gets more complicated. When these same variables go into a multivariate ordered logistic regression, urban-rural residence loses its significance entirely (OR = 0.978, $p = 0.50$). The urban-rural gap that looks enormous in a crude comparison essentially disappears once wealth, BMI and age are held constant. Urban women are not diabetic because they live in cities. They are diabetic because they tend to be older, wealthier and heavier than their rural counterparts. Strip out those compositional differences and residence tells you very little on its own.

This matters for policy. Targeting intervention by geography, placing clinics in cities and calling it done, misses the mechanism. The risk is not urban per se. It is a cluster of lifestyle variables that happen to be urban-coded in the current moment of India's transition.

Age Does the Heavy Lifting

No variable in this analysis performs like age. Among women aged 15 to 19, mean plasma glucose stood at 119.74 mg/dL. By the 45 to 49 age group it had reached 133.84 mg/dL. That monotonic increase across every five-year band was confirmed as highly significant by one-way ANOVA ($F = 56.17$, $p < 0.001$).

The regression results make the point even more forcefully. Compared to the youngest group, women aged 45 to 49 had odds nearly four times higher of being in a more severe glycaemic category (OR = 3.945, 95% CI: 3.57 to 4.36). The dose-response pattern across age groups is clean and consistent. Every additional decade adds meaningfully to risk.

Table 2: Prediabetes and Diabetes Prevalence by Age Group (Survey-Weighted Estimates)

Age Group	Normal (%)	Prediabetic (%)	Diabetic (%)
15-19	93.9	3.4	2.6
20-24	93.1	4.0	3.0
25-29	91.7	5.3	3.0
30-34	89.4	7.2	3.4
35-39	87.4	8.4	4.2
40-44	83.2	11.4	5.5
45-49	79.7	13.8	6.5

Source: NFHS-5; design-based F-test, p < 0.001

Look at that 45 to 49 row. One in five women in this age group is either prediabetic or diabetic. That is not a tail risk. That is a population-level emergency sitting one generation away from women who are currently in their twenties and thirties, gaining weight and moving less. The question is whether they will be reached in time to interrupt the trajectory.

When the age data is stratified by sex of household head, something noteworthy appears in its absence. The two strata track each other almost exactly. Women in female-headed households show marginally higher prevalence at every age band, but the differences are small and the age gradient completely dominates. Sex of household head, which carries significant weight in discussions of vulnerability and care access, does not appear here as a meaningful predictor of glycaemic status. The regression confirms this: OR = 1.047, p = 0.146.

BMI: The One That Works As Expected

Overweight and obese women had more than double the odds of being in a more severe diabetes category compared to underweight women (OR = 2.11, 95% CI: 1.95 to 2.28). Even women in the normal BMI range had 14 per cent higher odds than the underweight group. This is the finding that matches the global epidemiological story most closely, and it holds even after controlling for wealth, age, residence and other factors.

The BMI data also shows the clearest clinical gradient in prevalence terms. Among overweight and obese women, prediabetes stood at 11.3 per cent and diabetes at 5.1 per cent. Among underweight women those figures were 4.8 per cent and 1.3 per cent respectively. Adiposity and glucose dysregulation move together in this dataset in a way that other social variables do not. Age and BMI are the two predictors that survive full multivariate adjustment with the largest effect sizes. Everything else is either attenuated or explained by them.

Caste, Religion and the Residual

Caste differences in mean plasma glucose are significant in the bivariate analysis. Schedule Tribes had the lowest mean at 116.61 mg/dL. The Others category, broadly the general caste group, had the highest at 130.48 mg/dL. In the regression, this partially persists: ST women showed slightly lower odds of worse glycaemic status compared to scheduled caste women (OR = 0.908, p = 0.045). OBC and Others were not significantly different from the SC reference category once other factors were accounted for. The ST finding is marginal, but it does suggest that some part of the lower glucose burden in tribal communities reflects something beyond just age and BMI composition.

Religion is more striking. Muslim women had a mean plasma glucose of 139.22 mg/dL, the highest of any religious group, and a diabetes prevalence of 5.7 per cent against 3.4 per cent for Hindus. In the regression, the Muslim odds ratio remained at 1.237 (p < 0.001) after full adjustment. Christian women showed a similar pattern at OR = 1.170 (p = 0.029).

These effects are not explained away by wealth, age, BMI or residence. Something community-specific is operating here. Dietary patterns are the most obvious hypothesis. Geographic concentration in urban areas is a partial explanation for Muslim women but cannot account for the full residual effect. This is a finding worth investigating further with data that can disentangle household-level diet composition, physical activity patterns and healthcare utilisation by community.

The Education Puzzle

Education does not behave in the way one might expect. Women with no education had a mean plasma glucose of 126.22 mg/dL. Women with higher education had 132.57 mg/dL, the highest of any education group. Prediabetes prevalence was actually highest among women with no education (8.9 per cent) and primary education (9.0 per cent), while diabetes prevalence was highest among the higher educated (4.7 per cent).

The regression adds a further complication. Primary and secondary education showed slightly higher odds of worse diabetes severity compared to no education (OR = 1.19 and 1.12 respectively). Higher education was not significant at all (OR = 0.998, $p = 0.97$). This is almost certainly confounding. Educated women are disproportionately older, wealthier, urban and in higher BMI categories. Once those are controlled for, education's independent contribution dissolves.

What this tells us is that education's protective effect on health, which is well-documented in outcomes like maternal mortality and child immunisation, does not straightforwardly extend to glycaemic status in the current cross-section. The mechanisms may operate on a longer lag. Or education's effect may be primarily on health awareness and care-seeking, which shows up in earlier detection rather than lower prevalence.

What the Regression Settles

Ordered logistic regression was chosen as the modelling strategy because the dependent variable, classified as normal, prediabetic and diabetic, has a natural ordering that matters. Multinomial logistic regression would treat the three categories as unrelated. Binary logistic regression would collapse information. The proportional odds model preserves the structure of the outcome while providing a single odds ratio per predictor that applies across both thresholds.

The model fit well overall ($F(24, 29780) = 130.75$, $p < 0.001$, N = 188,766). Survey weighting was applied throughout using svyset and svy prefix commands to account for the complex sampling design of NFHS-5. Standard errors and significance tests are design-consistent.

Table 3: Key Predictors from Ordered Logistic Regression (Outcome: Diabetes Severity)

Variable	Odds Ratio	95% CI	p-value	Significant
Age 35-39 (ref: 15-19)	2.090	1.89 - 2.31	< 0.001	Yes
Age 40-44	3.015	2.73 - 3.33	< 0.001	Yes
Age 45-49	3.945	3.57 - 4.36	< 0.001	Yes
BMI >24.9 (ref: <18.5)	2.110	1.95 - 2.28	< 0.001	Yes
Religion: Muslim (ref: Hindu)	1.237	1.15 - 1.33	< 0.001	Yes
Religion: Christian	1.170	1.02 - 1.35	0.029	Yes
Residence: Rural (ref: Urban)	0.978	0.92 - 1.04	0.500	No

Variable	Odds Ratio	95% CI	p-value	Significant
Caste: ST (ref: SC)	0.908	0.83 - 1.00	0.045	Marginal
Wealth: Richest (ref: Poorest)	0.846	0.77 - 0.93	0.001	Yes

Source: NFHS-5; svy: ologit, N = 188,766, F(24, 29780) = 130.75, p < 0.001

The wealth reversal in the regression deserves emphasis. In crude analysis, wealth and diabetes track together positively: the richest quintile has the highest prevalence. In the adjusted model, every wealth group above the poorest shows significantly lower odds of worse glycaemic status compared to the poorest. This is not a contradiction. It is confounding. Wealthier women are older, heavier and more urban. Once those attributes are controlled, the apparent wealth advantage in crude terms flips into a protective residual. The poorest, holding constant all the risk factors that cluster in wealth, still face higher odds. This reflects something about healthcare access, quality of nutrition and chronic stress that wealth alone does not capture.

The Prediabetes Pipeline

The most urgent policy implication from this analysis is not about current diabetics. It is about the 7.1 per cent of women who are already prediabetic and largely unaware of it.

Among women aged 45 to 49, nearly 14 per cent are prediabetic. Among those who are overweight or obese, the figure is 11.3 per cent. These women are not yet diabetic. But without lifestyle intervention, a substantial proportion of them will be. The question is whether India's health system is structured to catch them at this stage.

The answer, at present, is probably not. Diabetes screening in primary care tends to target people who are already symptomatic or already diagnosed with a related condition. The infrastructure for population-level prediabetes detection, the community health worker capacity, the diagnostic tools, and the lifestyle counselling that has to follow a positive screen, is not where it needs to be. What this data makes visible is a group that is reachable today. The window to intervene will close.

Where to Go From Here

Three things stand out from this analysis as genuinely action-relevant rather than academically interesting.

First, the age-BMI combination is the dominant risk profile. Women over 35 with BMI above 24.9 carry dramatically elevated odds of glycaemic dysregulation. They are identifiable. They do not require elaborate biomarker screening to find. A community health worker with a weighing scale and a date of birth can flag them.

Second, the wealth reversal in adjusted analysis suggests that the poorest women are not protected. They appear protected in crude data because they are also younger, lighter and more rural. When those characteristics are held constant, poverty is associated with worse outcomes. Whatever is driving that residual, it likely includes poor diet quality, limited healthcare access and high stress load. Assuming that affluence is the only diabetes risk worth addressing is a mistake the data does not support.

Third, the Muslim and Christian religion effects persisting after adjustment warrant dedicated community-level research. Designing interventions that work culturally in communities where dietary habits, physical activity norms and healthcare trust patterns differ from the majority population is not a trivial problem. The first step is acknowledging that a residual community effect exists in this data after controlling for the obvious confounders.

India's non-communicable disease burden is growing. The diabetes transition is not coming. It is here. The NFHS-5 data shows clearly where it is concentrated, and more importantly, where it is about to expand. Acting on that picture, rather than waiting for hospitalisation records to tell the story a decade from now, is the practical imperative.

Data Source

International Institute for Population Sciences (IIPS) and ICF. National Family Health Survey (NFHS-5), 2019-21. India Report. Mumbai: IIPS, 2021. Analysis conducted in STATA using survey-weighted procedures (svyset, svy: tab, svy: ologit) on the NFHS-5 women's dataset.