CHAPTER 4

STUDY 1: USING A BETTER OBESITY MEASURE TO INCREASE SPECIFICITY OF IDRS

4.1 AIM AND OBJECTIVES

4.1.1 Aim of the study

To improve the classification characteristics of Indian Diabetes Risk score (IDRS).

4.1.2 Objectives of the Study

- To analyze the classic anthropometric measures of obesity, viz. Waist Circumference (WC), and Body Mass Index (BMI) in terms of their ability to encapsulate diabetes risk
- To define a composite measure of obesity (BMI_{WC}) that takes into account both BMI and WC, and analyze its ability to encapsulate diabetes risk as compared to BMI or WC in isolation
- To find the reliability of the new IDRS which uses the new measure of obesity (BMI_{WC})
- To compare the performance of IDRS with each of the three measures of obesity (BMI, WC, and BMI_{WC})

4.1.3 Rationale for the study

While obesity is an accepted risk factor for diabetes, the right anthropometric measure of obesity is unclear. Some studies support WC as the right measure, while others support BMI; and at least one study recommends both and at least one study recommends neither. Given that the evidence spans all four logical possibilities, this supports the theory that a composite measure that takes into account both general adiposity (as measured by BMI) and central fat

distribution (as measured by WC) would be a better than either one in isolation. Furthermore, since obesity is a component of almost all diabetes screening scores deployed in various countries, an improved measure of obesity has the potential to improve the efficacy of these screening scores.

4.1.4 Hypotheses and Null Hypotheses

Hypotheses

The present study hypothesized that (a) the composite measure BMI_{WC} is a better predictor of diabetes risk than either BMI or WC and (b) $IDRS_{BMIWC}$ has better screening characteristics than $IDRS_{WC}$ or $IDRS_{BMI.}$

NULL Hypotheses

- BMI_{WC} is not a better predictor of diabetes risk than BMI or WC
- $IDRS_{BMIWC}$ is not a better classifier of the population than $IDRS_{WC}$ or $IDRS_{BMI}$

4.2 DESIGN OF THE STUDY

Niyantrita Madhumeha Bharata ("Control of Diabetes in India") 2017, or NMB 2017, was a two-phased study undertaken across 29 most populous states/union territories in India. The twin objectives of the study were:

- (Phase 1) A rapid survey to estimate the prevalence of diabetes, prediabetes, and high-risk population using IDRS simultaneously in all zones of India in 2017
- (Phase 2) To conduct an RCT using a validated yoga life-style protocol

Phase 1 (Nagendra et al., 2019) was a nationwide randomized cross-sectional survey using a multi-level stratified cluster sampling technique with random selection among urban and rural populations covering 65 districts of the most populous states (25) and union territories (4) of

the country. In a door to door survey, researchers used short a questionnaire to collect data on diabetes status and diabetes risk.

Phase 2 (Nagendra et al., 2019) involved a sub-sample of high-risk individuals from the participants of phase-I; these were (a) those with high risk for diabetes (IDRS $\geq$ 60) and (b) all those with self-reported diabetes (with or without high risk). Phase 2 carried out further assessments through blood tests and a more detailed questionnaire and implemented a 2 armed controlled cluster randomized yoga based life style intervention to determine its efficacy . The intervention was a 3-month practice of a standard Yoga protocol (Nagarathna, et al., 2019).

4.2.1 Phase 1 Sampling Strategy

Sampling was done at 4 levels: (a) Zones (b) States (c) Randomly selected Districts, and (d) Randomized cluster sampling of Villages (rural) or Towns (urban), with a further randomized cluster sampling of Census Enumeration Blocks (CEBs) in each town. We chose 24 (of 29) states and 5 (of 7) union territories. These were grouped into 7 geographical regions based on their sociocultural similarities with a small deviation from the grouping available in the national directory ("Administrative divisions of India - Wikipedia"). Table 4 shows the composition of each zone.

Table 4: States and Union Territories in each zone for NMB-2017

Zone	States/Union Territories
North-West	Jammu and Kashmir + Ladakh
North-East	Arunachal Pradesh, Assam, Manipur, Tripura, Meghalaya
North	Delhi, Punjab, Chandigarh, Haryana, Uttarakhand, Uttar Pradesh, Himachal Pradesh
West	Rajasthan, Gujrat, Maharashtra
Central	Madhya Pradesh, Jharkhand
East	Bihar, Chhattisgarh, West Bengal, Odisha
South	Karnataka, Goa, Kerala, Tamil Nadu, Andhra Pradesh + Telangana, Pondicherry, Andaman and Nicobar

The districts in the state were the first level of sampling. In accordance with the sampling plan, it was decided to select 10% of the total number of districts in the country, and correspondingly 10% of the districts in each state. For states with ≤ 10 districts, one district was chosen, and for states with 10 – 30 districts, two districts were chosen. To ensure that the district samples within a state were not clustered, we grouped the state into geographical regions and chose a district from each region (e.g., if a state needed 3 districts, it was grouped into north, south, and central).

A statistician selected two districts randomly within the group (i.e. double the number needed); after a review of local conditions within the chosen district, one of the two districts were purposively sampled.

Each district was also grouped into four geographical regions (north, east, south, and west).

- (Rural) Within each region, a statistician chose at random two villages from among villages with a population of around 500. After a review of local conditions, one of the two villages were purposively sampled. This resulted in four villages with a population of around 2000.
- (Urban) Within each region, a statistician chose one town/city. From this list of four towns/cities, one city was purposively chosen after a review of local conditions. Within the chosen town/city, a ward was randomly selected. The selected ward was grouped into four geographical regions (north, east, south, and west). Depending on the size of the CEBs (which are either 500 people or 1000 people), either two or four wards were randomly selected ensuring that each CEB came from a different geographical region.

All individuals ≥ 20 years of age in each household of the selected CEB were screened.

4.2.2 Phase 2 Sampling Strategy

From the Phase 1 sample, we selected adults of both genders who had the ability to do yoga (and consented to doing it), and satisfied one of the following criteria:

1. Self-reported and newly diagnosed diabetes with or without glycemic control, with or without high risk scores, using/not using oral hypoglycemic agents or insulin
2. Did not self-report diabetes, but IDRS score was ≥ 60

4.2.3 Data Collection

Two research associates, 7 zonal coordinators, 35 senior research fellows, and 1200 data-collectors were involved in data collection. Each group was given training on their respective duties.

Phase 1 questionnaire (Nagendra et al., 2019) included the following data:

1. Demographic: age, gender, marital status, education level, socio-economic status, and occupation
2. Anthropometric: height, weight, hip/waist circumference
3. Diabetes risk: family history, physical activity, prior diagnosis of diabetes (and for how long)
4. Two blood pressure readings at 5 min interval.
5. Yoga practice if any, and details of the practice.

Phase 2 questionnaire (Nagarathna, et al., 2019) included (in addition to Phase 1 data) the following data:

1. Demographic: personal income, number of earning members in the family, type of work, etc.

2. Detailed medical history, including diabetes, CVD, etc.

3. Stress level assessment

4. Lifestyle: tobacco/alcohol use, diet, sleep

Phase 2 individuals also underwent blood tests which measured FBS, PPBS, MBG, HbA1c, and Lipid profile

4.2.4 Procedure for anthropometric and biochemical measures

Anthropometric measurements were conducted by measuring weight in kg, height and waist circumference in cm.

All biochemical assays were carried out by the same method by the same laboratory that has a nationwide representation (SRS Labs) during the study period. HbA1c was estimated by high-pressure liquid chromatography using Variant^{TM} II Turbo (Bio Rad, Hercules, CA). This is certified by the National Glyco-hemoglobin Standardization Program as having documented traceability to the Diabetes Control and Complications Study reference method. Serum cholesterol (cholesterol esterase oxidase-peroxidase-amidopyrine method), serum triglycerides (glycerol phosphate oxidase-peroxidase-amidopyrine method), and high-density lipoprotein cholesterol (direct method; polyethylene glycol-penetrated enzymes) were measured using an autoanalyzer (Model 2700/480; Beckman Coulter AU, Olympus, County Clare, Ireland). Serum creatinine was measured using the Jaffe method. The intra- and inter-assay coefficients of variation for the biochemical assays ranged from target goals for HbA1c, BP, and LDL measures were based on ADA's Standards of Medical Care for most people with diabetes including HbA1c < 7% (<53 mmol/mol), BP < 130/80 mmHg, and LDL < 100 mg/dL. Glycemic targets were the primary targets of the study.

4.3 METHODS

4.3.1 Participants and outcomes

We included only those individuals of Phase 2 for whom all the following data were available: WC, Weight, Height, Family history of diabetes, Age, Physical Activity, HbA1c, and Diabetes Self Declaration (Yes or No).

4.3.2 Sample Size

The sample size was calculated for the primary outcome for the study: relative risk reduction (RRR), which is relative decrease in risk of conversion to diabetes between the yoga and control groups within the study period of 3 months. We used RRR reported in individuals with prediabetes in the *Diabetes Community Lifestyle Improvement Program study (D-CLIP)* on 578 overweight/obese Asian Indian adults (Weber, Ranjani, Meyers, Mohan, & Narayan, 2012). The authors reported annual incidence rates diabetes as 11.1% and 7.8% in the control and intervention conditions, respectively. This provided a conversion rate at a 3-month follow-up to be P1 = 2.76% at control condition, and P2 = 1.95% at intervention condition. The following formula was used to calculate the sample size N (Hajian-Tilaki, 2011):

$$N = \frac{2P(1-P)(Z_{\frac{\alpha}{2}} + Z_{\beta})^2}{(P1-P2)^2}$$

Where P1 = probability of T2DM incidence in control condition (2.76%), P2 = probability of incidence in intervention condition (1.95%), α and β are the statistical significance (set to 0.05) and power (set to 0.8) respectively, $Z_{\frac{\alpha}{2}}$ and Z_{β} are the corresponding Z-scores, and $P = \frac{(P1+P2)}{2}$.

This gave us a sample size of 5314 for each group, for a total of 10,628 participants. Factoring an attrition of 20%, the final sample size was estimated to be 12,754 individuals with prediabetes. The best estimate of the prevalence of prediabetes in India is 5.5%, leading to the

requirement to screen 231,883 adults above the age of 20 years. It was determined that the population to be screened would be derived from 60 districts, which represented 10% of all districts in India (according to 2011 Census of India). Consequently, the study targeted approximately 4000 adults per district, with equal distribution among urban and rural areas.

4.3.3 Definitions of Obesity Metrics

Values of WC and BMI were bucketed into five risk categories (Table 5). The 5 categories for BMI were picked from the standardized ranges established for Asian populations (Nishida et al., 2004). For WC, we added two more categories at the bottom and top of the three categories established for the Asian Indian population (Mohan et al., 2005).

We created a composite obesity metric, BMI_{WC}, which combines BMI and WC according to the following algorithm: If WC was < 3, then BMI was scored as BMI – 1; if WC was ≥3, the value of BMI remained unchanged. Thus, BMI_{WC} recognizes that individuals with both low WC and low BMI are at lower risk while individuals with either high WC or high BMI are at higher risk. This adds an additional risk category at the lower end, with a score of zero, designated "Ultra Low" (Table 5).

Table 5 – Definitions of Obesity Metrics

Metric	**Risk Score**
WC Value (in cm)	
≤ 69·99 * (female), ≤ 79·99 * (male)	1 =Very Low (VL)
70 to 79·99 (female), 80 to 89·99 (male)	2 =Low (L)
80 to 89·99 (female), 90 to 99·99 (male)	3 =Moderate (M)
90 to 99·99 (female), 100 to 109·99 (male)	4 =High (H)
≥ 100 * (female), ≥110* (male)	5 =Very High (VH)
BMI Value (in kg/m^2)	
≤ 18·49	1 =Very Low (VL)
18·5 to 22·99	2 =Low (L)
23 to 27·49	3 =Moderate (M)
27·5 to 32·49	4 =High (H)
≥32·5	5 =Very High (VH)
BMI_{WC} (dimensionless), values of BMI and WC below refer to risk scores	
BMI=1 & WC<3	0 = Ultra Low (UL)
BMI=2 & WC<3 OR BMI=1 & WC≥ 3	1 =Very Low (VL)
BMI=3 & WC<3 OR BMI=2 & WC≥ 3	2 =Low (L)
BMI=4 & WC<3 OR BMI=3 & WC≥ 3	3 =Moderate (M)

BMI=5 & WC<3 OR BMI=4 & WC≥ 3	4 =High (H)
BMI=5 & WC≥ 3	5 =Very High (VH)
* Two additional categories added at the top and bottom of the three categories established for Asian Indian populations	

Below are some examples of obesity risk scores, calculated using data from NMB 2017:

- Male with WC 85cm, BMI 27·6kg/m^2 has: WC=2, BMI=4, and BMI_{WC}=3

- Male with WC 108cm, BMI 26.3kg/m^2 has: WC=4, BMI=3, and BMI_{WC}=3

4.3.4 Definitions of IDRS and its variants

The second part of the study sought to validate the efficacy of BMI_{WC} by replacing the obesity component of IDRS (WC) with BMI_{WC}. We also studied the efficacy of IDRS when the obesity component is replaced by BMI. The modified risk scores were called $IDRS_{BMIWC}$ and $IDRS_{BMI}$ resp. Table 6 shows the definitions of $IDRS_{WC}$ (Mohan et al., 2005), $IDRS_{BMI}$ and $IDRS_{BMIWC}$.

Table 6 – Definitions $IDRS_{WC}$, $IDRS_{BMI}$, and $IDRS_{BMIWC}$

Metric	Score
$IDRS_{WC}$	
Age	
< 35 years	0
35-49 years	20
≥ 50	30
Physical Activity	
Exercise [regular] + strenuous work	0
Exercise [regular] or strenuous work	20
No exercise and sedentary work	30
Family History	
No family history	0
Either parent	10
Both parents	20
Obesity (WC)	
WC Risk Score ≤ 2	0
WC Risk Score = 3	10
WC Risk Score ≥ 4	20
Range of the Score	0-100
$IDRS_{BMI}$	
Age, Physical Activity, Family History are same as IDRS	0-80
Obesity (BMI)	
BMI Risk Score ≤ 2	0
BMI Risk Score = 3	10
BMI Risk Score ≥ 4	20
Range of the score	0 – 100
$IDRS_{BMIWC}$	
Age, Physical Activity, Family History are same as IDRS	0-80
Obesity (Composite)	

If WC Risk Score ≤2	0
BMI Risk Score ≤ 2	0
BMI Risk Score = 3	10
BMI Risk Score ≥ 4	20
If Waist Risk Score > 2	
BMI Risk Score ≤ 2	10
BMI Risk Score = 3	20
BMI Risk Score ≥ 4	30
Range of the score	0-110

4.3.5 Analysis

Contingency table methods (for risk assessment) and confusion matrix methods (for assessing classification efficacy) were used to evaluate each obesity metric. Validation was done by replacing WC with BMI_{WC} as the obesity component of IDRS and determining classification efficiency of the modified IDRS.

WC, BMI, and BMI_{WC} were compared for their association to type 2 diabetes risk. A contingency table of risk categories and outcome was created for each metric, and χ^2 statistic was calculated to measure risk association. Using the lowest risk category as a reference, Odd Ratio (OR) calculated for each risk category. They were also compared for their ability to classify the population into two groups: people with type 2 diabetes and people without. An ROC curve was drawn for each measure to determine the threshold score for classification. Based on this threshold, a confusion matrix was created for each measure. Efficacy of classification was determined by calculating Sensitivity, Specificity, and Accuracy (Pepe, 2004). McNemar's statistic was calculated to determine the statistical significance of the difference in Specificities (Hawass, 1997).

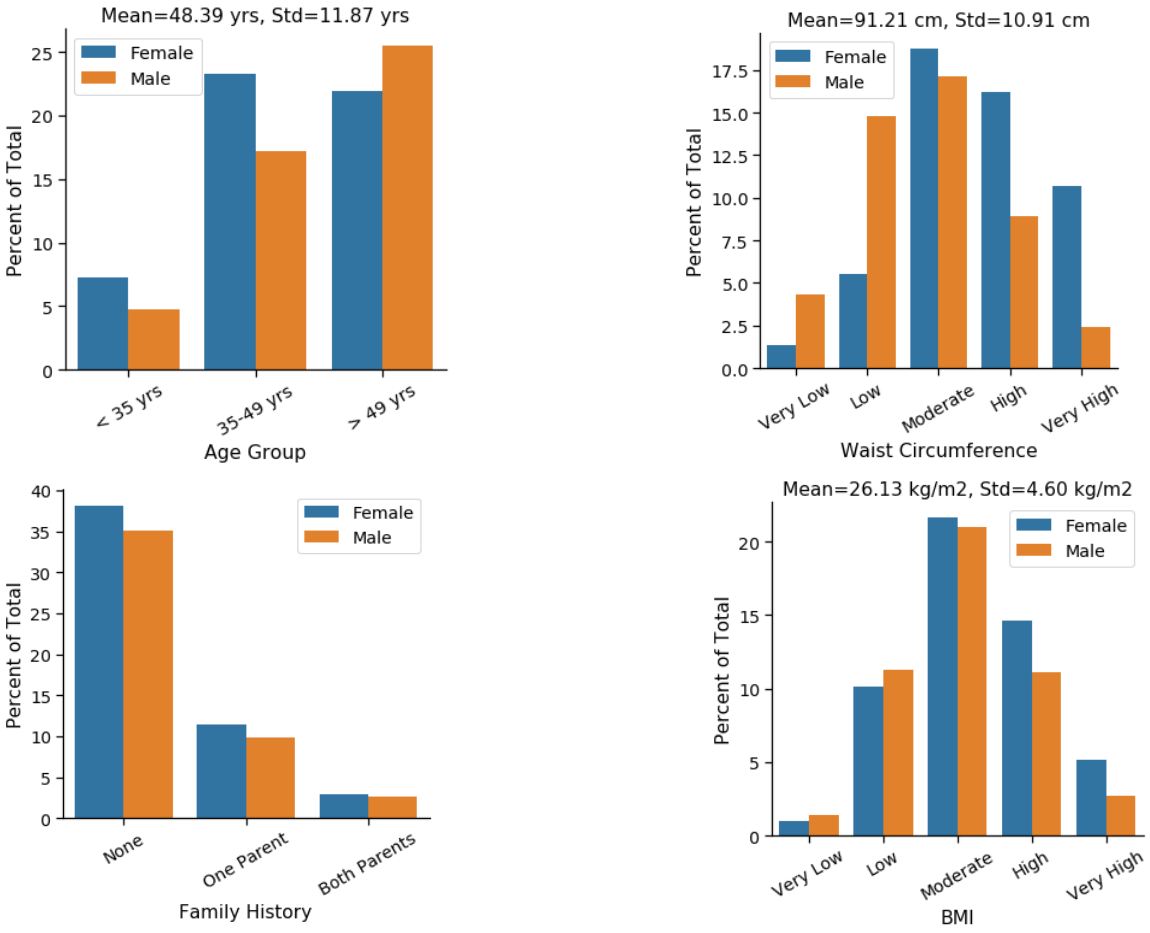
$IDRS_{WC}$, $IDRS_{BMI}$ and $IDRS_{BMIWC}$ were compared for efficacy of classification. An ROC curve was drawn for $IDRS_{BMI}$ and $IDRS_{BMIWC}$ to determine classification thresholds. The threshold for IDRS has already been determined to be 60 (Mohan et al., 2005). Using these threshold values, Sensitivity, Specificity and Accuracy were calculated. McNemar's statistic was calculated to as before to determine statistical significance. All analyses were done using

Python v.3·7. Pandas v.0·23 was used to import data, calculate obesity metrics and risk levels. Contingency table creation and calculation of risk measures were done using Statsmodels v.0·10·1. Confusion matrix creation and calculation of classification measures were done using Scikit-learn v.0·21·3. χ^2 and McNemar's statistics were calculated using Scipy v.1·3·0.

4.4 RESULTS

4.4.1 Description of Data

Fig 7 shows the distribution of each of these characteristics across relevant categories.



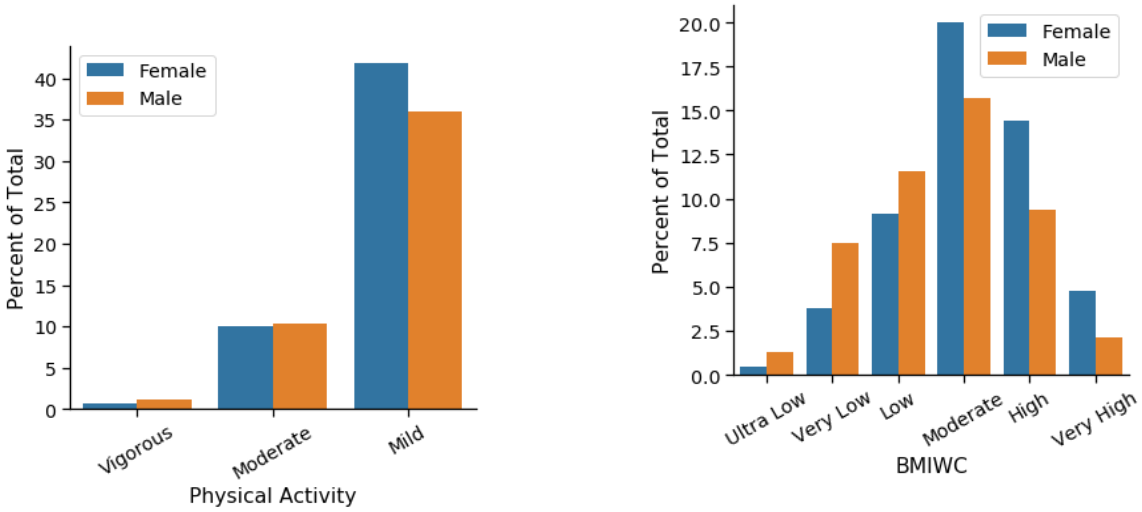


Fig 7: Respondent Characteristics, n=7496

A total of 7496 individuals at high risk (≥60 on IDRS) for type 2 diabetes (3935 females, 3561 males) were analyzed. They varied in age from 20 - 85 years (μ = 48·39, σ = 11·86). Waist circumference varied from 60 - 150 cm (μ = 91.21, σ = 10.91) and BMI varied from 12.2 – 66.2 kg/m^2 (μ = 28.13, σ = 4.60).

Total number with type 2 diabetes was 3079, of which 1093 individuals were newly diagnosed and 1986 were self-reported.

4.4.2 Risk analysis of obesity metrics

The χ^2 test of association showed statistically significant association between obesity metrics and type 2 diabetes risk: WC: $\chi^2(4, N = 7496) = 29{\cdot}10$, $p < 0{\cdot}001$; BMI: $\chi^2(4, N = 7496) = 66.58$, $p < 0{\cdot}001$; BMI_{WC}: $\chi^2(5, N = 7496) = 59.06$, $p < 0{\cdot}001$.

Odds that a person in the lowest obesity category (VL for WC and BMI, UL for BMI_{WC}) had diabetes was calculated for each obesity metric, which was used as the reference odds. Odds were also calculated at each of the higher obesity categories, and the odds ratio was determined by taking the ratio of this with the reference odds.

The following OR values were seen at the highest risk category (VH) for each obesity metric:

- For WC: 1·87 (95% CI 1.47 – 2.37)
- For BMI: 2·26 (95% CI 1.58 – 3.24)
- For BMI_{WC}: 2·30 (95% CI 1.51 – 3.51)

We can see that WC, BMI and BMI_{WC} each have higher odds in the VH category compared to the reference (lowest) category. But BMI_{WC} outperformed WC and BMI by having a higher OR.

We also observed that the OR for BMI_{WC} was higher at every risk category than the corresponding scores for WC and BMI, as seen in Fig 8.

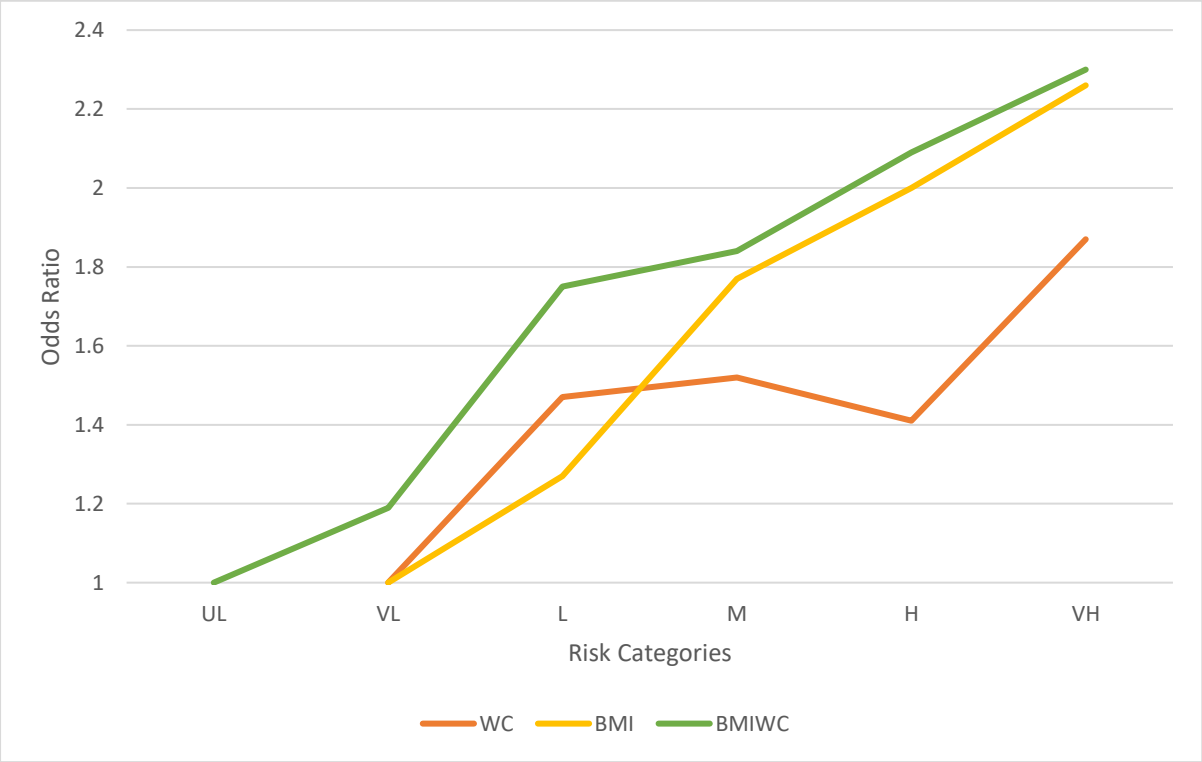


Fig 8: Odds Ratio for WC, BMI, and BMI_{WC}

WC showed an actual decrease in OR between Moderate (M) and High (H) risk levels but showed a dramatically increased odds between High (H) and Very High (VH). This non-monotonic behavior is an indication that the risk categories of WC don't adequately capture increasing diabetes risk. BMI encapsulates diabetes risk better by showing a monotonically

increasing OR. But BMI_{WC} clearly outperforms the WC and BMI: OR is monotonically increasing, and the value of OR is higher at every risk category – as can be seen by the green line (representing BMI_{WC}) lying above the orange (WC) and yellow (BMI) lines.

4.4.3 Classification analysis of obesity metrics

We plotted ROC curves for WC, BMI, and BMI_{WC} to determine the classification thresholds for each measure. These curves are shown in Fig 9.

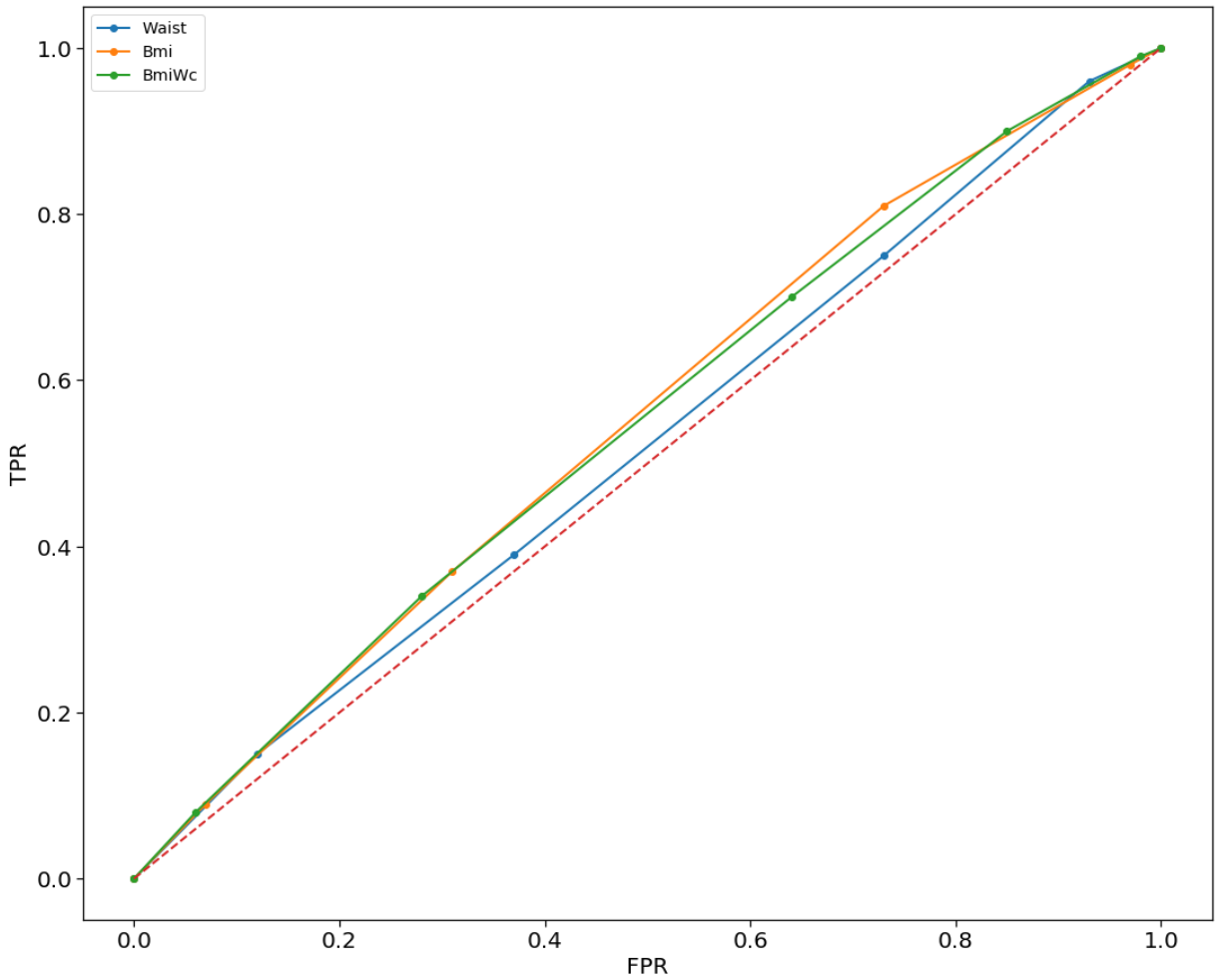


Fig 9. ROC Curves for WC, BMI, and BMI_{WC}

We can see that a risk level of three (Moderate) is the optimum threshold, since the y-intercept is the highest for this risk level each of the measures: WC (blue), BMI (orange), and BMI_{WC} (green) . Using this value, we calculated Sensitivity, Specificity, and Accuracy. Table 7 shows

the results. As we can see, BMI had better Sensitivity (6·89%) when compared to WC but showed the same Specificity. BMI_{WC} showed slightly decreased Sensitivity (-7·00%) but vastly improved Specificity (34·01%) when compared to WC. In terms of Accuracy, BMI was slightly better than WC (4·41%), and BMI_{WC} was better still (6·69%).

Table 7: Results of classification analysis

Metric	Sensitivity	Specificity	Accuracy
WC	0·75	0·27	0·47
BMI	0·81	0·27	0·49
BMI_{WC}	0·70	0·36	0·50
$IDRS_{WC}$	0·87	0·30	0·53
$IDRS_{BMI}$	0·88	0·31	0·54
$IDRS_{BMIWC}$	0·82	0·39	0·56

Matched sample tables for Specificity were created using True Negative (TN) and False Positive (FP) counts, one for BMI_{WC} and WC, and another for BMI_{WC} and BMI. Table 8 shows the counts of tied (FP-FP, TN-TN) and untied (TN-FP, FP-TN) pairs. McNemar's statistic calculated on the values untied pairs in these tables (Hawass, 1997). The results were: BMI_{WC} and WC: χ^2 (1, N = 695) = 234·84, p < 0·001; BMI_{WC} and BMI: χ^2 (1, N = 410) = 408·00, p < 0·001. This shows that the increase the Specificity of BMI_{WC} as compared to WC and BMI is statistically significant.

Table 8: Matched samples tables for Specificity

BMI_{WC} / WC	FP	TN
FP	2676	145
TN	550	1046

BMI_{WC} / BMI	FP	TN
FP	2821	0
TN	410	1186

$IDRS_{BMIWC}$ / $IDRS_{WC}$	FP	TN
FP	2605	104
TN	463	1245

$IDRS_{BMIWC}$ / $IDRS_{BMI}$	FP	TN
FP	2709	0
TN	334	1374

4.4.4 Classification analysis of IDRS variants

We plotted ROC curves for $IDRS_{BMI}$, and $IDRS_{BMIWC}$ to determine the classification thresholds for each score. These curves are shown in Fig 10. We can see that 60 is the optimum threshold for $IDRS_{BMI}$ (the orange line), and 70 is the threshold for $IDRS_{BMIWC}$ (represented by the green

line). This is based on the maximal y-intercept value. The threshold for IDRS has already determined to be 60 (Mohan et al., 2005). These values were used to calculate Sensitivity, Specificity, and Accuracy. Table 7 shows the results.

$IDRS_{BMI}$ showed marginally better Sensitivity (1·27%) and Specificity (1·85%) when compared to $IDRS_{WC}$. $IDRS_{BMIWC}$ showed slightly decreased Sensitivity (-6·14%) but vastly improved Specificity (26·61%) when compared to $IDRS_{WC}$. In terms of Accuracy, $IDRS_{BMI}$ was slightly better than $IDRS_{WC}$ (1·46%), and $IDRS_{BMIWC}$ was better still (4·79%).

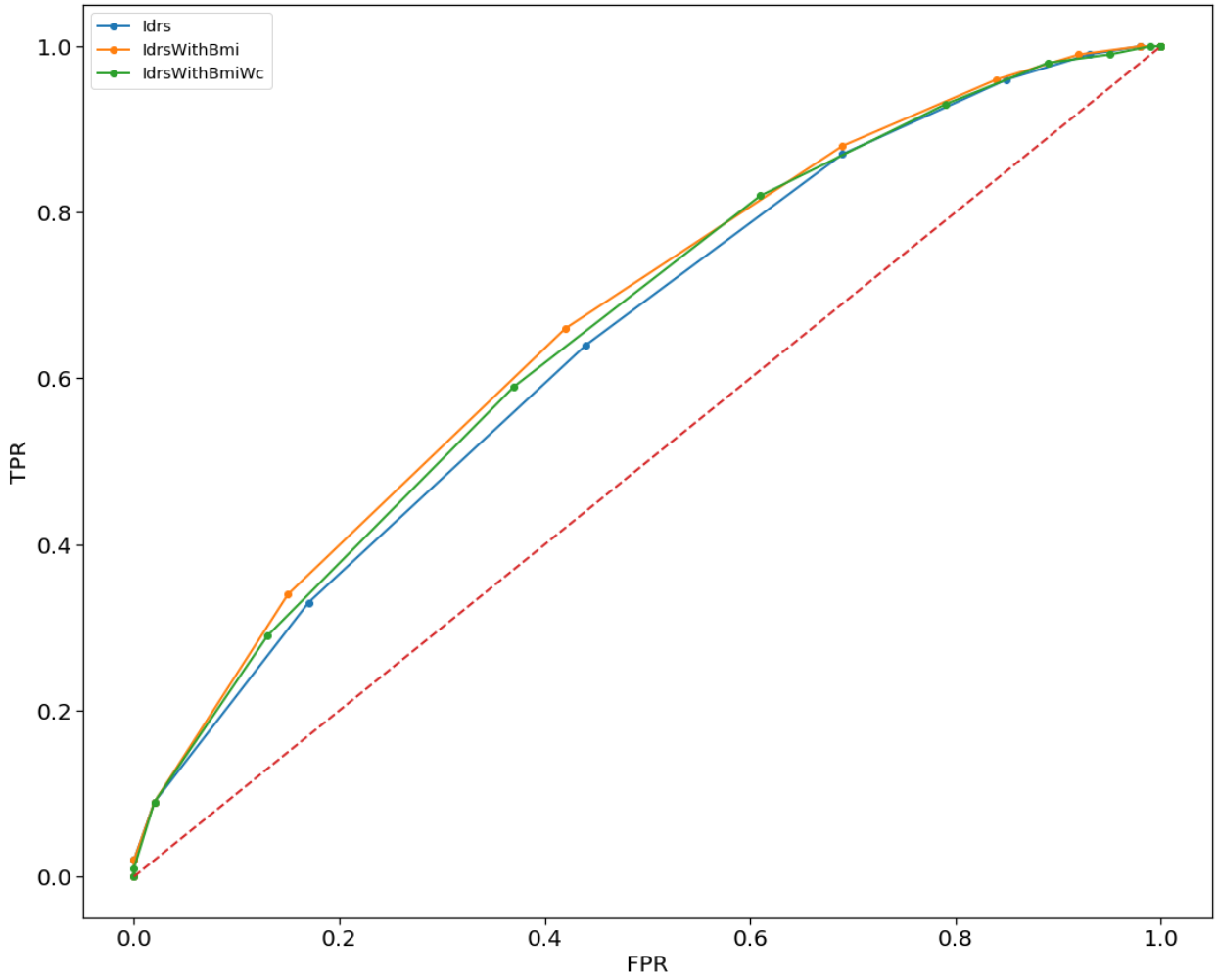


Fig 10. ROC Curves for IDRS, $IDRS_{BMI}$, and $IDRS_{BMIWC}$

Matched sample tables for Specificity were created using True Negative (TN) and False Positive (FP) counts, one for $IDRS_{BMIWC}$ and $IDRS_{WC}$, and another for $IDRS_{BMIWC}$ and

$IDRS_{BMI}$. Table 8 shows the counts of tied (FP-FP, TN-TN) and untied (TN-FP, FP-TN) pairs. McNemar's statistic calculated on the values untied pairs in these tables (Hawass, 1997). The results were: $IDRS_{BMIWC}$ and IDRS: χ^2 (1, N = 567) = 226·04, p < 0·001; $IDRS_{BMIWC}$ and $IDRS_{BMI}$: χ^2 (1, N = 334) = 332·00, p < 0·001. This shows that the increase the Specificity of $IDRS_{BMIWC}$ as compared to $IDRS_{WC}$ and $IDRS_{BMI}$ is statistically significant.