medRxiv preprint doi: https://doi.org/10.1101/19007872.this version posted October 5, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Predicting youth diabetes risk using NHANES data and machine learning

Running title: Youth diabetes risk screening

Nita Vangeepuram[1,2,3,*], Bian Liu[2,3], Po-hsiang Chiu[4], Linhua Wang[4,X], Gaurav Pandey[4]

[1] Department of Pediatrics, Division of General Pediatrics, Icahn School of Medicine at Mount Sinai, New York, NY, USA
[2] Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, New York, NY, USA
[3] Department of Environmental Medicine & Public Health, Icahn School of Medicine at Mount Sinai, New York, NY, USA
[4] Department of Genetics and Genomic Sciences and Icahn Institute for Data Science and Genomic Technology, Icahn School of Medicine at Mount Sinai, New York, NY, USA
[x] Current address: Baylor College of Medicine, Houston, TX, USA
[*] Corresponding author, E-mail: nita.vangeepuram@mssm.edu

Acknowledgements: The research presented in this manuscript was supported by a National Institutes of Health grant (R01GM114434) and an IBM Faculty award to author GP and by a Cigna Foundation grant (10005177) awarded to author NV. The study was enabled in part by computational resources provided by Scientific Computing at the Icahn School of Medicine at Mount Sinai.

Author Contributions: N.V., B.L., and G.P. conceived the study and wrote the manuscript. N.V. provided clinical expertise and supervised the study. B.L. prepared the relevant NHANES data and carried out the performance analyses of the screeners. L.W. and P.C. carried out the machine learning analyses under G.P.'s supervision. All the authors reviewed and approved the manuscript.

Abstract

Type 2 diabetes has become alarmingly prevalent among youth in recent years. However, simple questionnaire-based screening tools to reliably identify diabetes risk and prevent the adverse effects of this serious disease are only available for adults, not for youth. As a first step in developing such a tool, we used a large-scale dataset from the National Health and Nutritional Examination Survey (NHANES), to examine the performance of a well-known adult diabetes risk self-assessment screener and published pediatric clinical screening guidelines in identifying youth with pre-diabetes/diabetes (pre-DM/DM) based on American Diabetes Association diagnostic biomarkers. We assessed the agreement between the adult screener/pediatric screening guidelines and biomarker diagnostic criteria by conducting comparisons using the overall data set and sub-datasets stratified by sex, race/ethnicity, and age. While the pediatric guidelines performed better than the adult screener in identifying youth with pre-DM/DM (sensitivity 43.1% vs 7.2%), both are inadequate for general deployment among youth. There were also notable differences in the performance of the pediatric guidelines across subgroups based on age, sex and race/ethnicity. In an effort to improve pre-DM/DM screening, we also evaluated data-driven machine learning-based

medRxiv preprint doi: https://doi.org/10.1101/19007872.this version posted October 5, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

classification algorithms, several of which performed slightly but statistically significantly better than the pediatric screening guidelines.

Key words: Adolescent; Diabetes Mellitus, Type 2; Surveys and Questionnaires; Biomarkers; Machine Learning

Introduction

Diabetes mellitus (DM) is a serious chronic condition associated with numerous long-term complications, including cardiovascular disease, neuropathy, kidney disease, eye disorders, poor wound healing, hearing impairment, and Alzheimer's disease.[1] The total estimated cost of diabetes in the U.S. in 2017 was $327 billion, including direct medical costs and those associated with reduced productivity.[2] Pre-diabetes (pre-DM) is a precursor condition in which glucose levels are high, but not yet high enough to diagnose diabetes.[3] Unlike DM, pre-DM is reversible with lifestyle modification and weight loss, offering an avenue to avoid the adverse effects of diabetes [3-5] Both these conditions have become alarmingly prevalent among youth in recent years.[6,7] According to a large prospective cohort study, an estimated 5,300 youth age 10 to 19 are newly diagnosed with type 2 DM annually in the US,[6] with a higher prevalence among older teens than younger ones.[7] The overall prevalence of pre-DM among US adolescents based on nationally representative data from the National Health and Nutritional Examination Survey (NHANES) was 17.7%, with higher rates in males (22.0%) than females (13.2%).[8] Pre-DM prevalence was also higher in non-Hispanic Blacks (21.0%) and Hispanics (22.9%) than in non-Hispanic Whites (15.1%), highlighting a racial disparity challenge as well.[8] Compared to adults, DM diagnosed in youth is more difficult to treat [9] due to a more rapidly progressive decline in beta cell function, and an earlier onset of complications.[10,11] The potential health and economic impact of DM is therefore even greater for youth than adults, given the greater number of years living with the disease and time to develop long-term complications.

The American Diabetes Association (ADA) has published guidelines for the diagnosis of pre-DM and DM among youth.[12] A two hour oral glucose tolerance test (OGTT), including measurement of plasma glucose levels after an overnight fast (FPG), and again two hours after a 75 gram oral glucose load (2hrPG), is used to diagnose pre-DM and DM based on predefined cutoffs.[12] DM is diagnosed when FPG$\geq$126 mg/dL and/or 2hrPG$\geq$200 mg/dL, while pre-DM is diagnosed when FPG is 100-125 mg/dL (impaired fasting glucose) and/or 2hrPG is 140-199 mg/dL (impaired glucose tolerance). In 2010, the ADA expanded diagnostic guidelines to three biomarkers, namely FPG, 2hrPG, and hemoglobin A1c (HbA1c).[12] Although the ADA acknowledges the limited data supporting use of HbA1c alone for diagnosing pre-DM and DM in children and adolescents,[12] the association highlights several reasons to consider its use. These include wide availability, ability to perform the test at any time of the day, increased provider adherence to screening recommendations compared to OGTT, and lesser day-to-day variations due to factors like stress or illness.[5,13]

In spite of these diagnostic guidelines, pre-DM and DM are often underdiagnosed among youth, as many affected individuals have no or few symptoms.[14] For example, one study found that only 1% of adolescents with pre-DM based on elevated HbA1c or fasting glucose reported having been told by a physician that they had the condition.[15]

medRxiv preprint doi: https://doi.org/10.1101/19007872.this version posted October 5, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

In addition, despite professional consensus on recommended preventive care for adolescents, many youth do not receive annual checkups, and receipt of preventive services remains suboptimal.[16,17] Even for those in care, OGTT is not conducted by most general pediatricians.[18,19] Furthermore, OGTT requires fasting and testing over 2-3 hours, which is often challenging,[20,21] especially in communities most at risk for pre-DM/DM.[8] Thus, many youth with pre-DM/DM may be unaware of their condition, making it difficult to target the highest risk youth for prevention. A simple non-invasive, questionnaire-based screening tool is therefore a likely impactful first-line strategy to identify at-risk individuals before subjecting them to definitive testing and resource-intense prevention programs.[22-25]

Several such risk assessment tools have been developed to detect the risk of prevalent (undiagnosed) and incident pre-DM and DM in adults.[26-33] For example, the ADA and the Centers for Disease Control (CDC) have developed an easy-to-use, cost-effective patient self-assessment screener based on 7 questions to identify adults at risk for pre-DM and DM.[34,35] Those determined to be at risk can then follow up with a physician for formal clinical diagnosis with an OGTT and/or hemoglobin A1c test.[12] Surprisingly, despite the clinical and public health importance of accurately screening for pre-DM/DM risk among youth, there exists no similar readily available screening tool for this purpose. The most widely used clinical screening guidelines for health care providers to test asymptomatic children and adolescents were published by the ADA and endorsed by the American Academy of Pediatrics (AAP).[12] These pediatric clinical screening guidelines state that children over 10 years of age, or those who have entered puberty (even if younger than 10 years), should be considered at high risk for pre-DM/DM if they are overweight/obese (BMI ≥ 85th percentile) and have two or more additional risk factors (**Table 1**).[12] However, these guidelines have not been validated using large population health data sets and ADA diagnostic guidelines in youth.[12] Furthermore, such screening guidelines may not perform equally in different age, sex and race/ethnicity subgroups.[36]

To address these gaps in our knowledge, we used a large-scale dataset from NHANES[37], to: 1) examine the performance of the adult ADA/CDC patient self-assessment screener in identifying youth with pre-DM/DM based on ADA diagnostic guidelines (elevated FPG, 2hrPG, or HbA1c), and 2) examine the performance of the AAP/ADA pediatric clinical screening guidelines in identifying pre-diabetic youth based on diagnostic guidelines, and examine how these screening guidelines perform in age, sex, and racial/ethnic subgroups. While NHANES data have been used to assess the prevalence of pre-DM and DM in both adults and youth,[8,38,39] no studies have used NHANES data to test the performance of existing pre-DM/DM pediatric clinical screening guidelines. Furthermore, recent work in various biomedical and other domains have shown that hypothesis-free data-driven machine learning (ML) methods [40,41] can help develop predictive models that can improve disease diagnosis, prognosis, and treatment efficacy.[42-47] Inspired by these advances, we also investigated if ML methods applied to NHANES data can help improve pre-DM/DM screening performance.

Methods

<u>Study population</u>

medRxiv preprint doi: https://doi.org/10.1101/19007872.this version posted October 5, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

We utilized publicly available data from NHANES, an ongoing cross-sectional national survey that systematically gathers data from interviews, physical examinations, and laboratory testing, and is used to study a range of health topics.[37,48] Using a complex, multi-stage probability cluster sampling design, this survey reflects a nationally representative sample of the US non-institutionalized population. NHANES also oversamples certain subgroups, such as African-Americans, Hispanics, Asians, older adults, and low income populations, to achieve reliable and precise estimates of health status indicators for these groups. Since 1999, NHANES data have been released to the public on a bi-annual cycle, with approximately 10,000 new people in each survey wave.

The sample used in this study consists of youth aged 12 to 19 years at the time of being surveyed. We combined NHANES data from 6 survey cycles for which pre-DM/DM biomarker data were available: 2005-2006, 2007-2008, 2009-2010, 2011-2012, 2013-2014, and 2015-2016.[49] The survey response rates for each cycle for individuals ages 12 through 19 years were approximately 82.1%, 83.6%, 85.0%, 76.1%, 76.5%, and 64.3%, respectively. We only included participants for whom pre-DM/DM status could be determined based on FPG, 2hrPG, and HbA1c (n=2970). We excluded data from 112 participants that lacked information on BMI percentile, family history of diabetes, blood pressure measures or total cholesterol, thus making it impossible to apply the AAP/ADA screening guidelines, resulting in a final sample size of 2858.

Diagnostic Tests Used to Classify Pre-diabetes/Diabetes Status

We identified youth from NHANES with pre-DM or DM based on current ADA diagnostic guidelines (elevated levels in any of the three biomarkers: FPG ≥100 mg/dL, 2hrPG ≥140 mg/dL, or HbA1C≥ 5.7%).[12] Serum specimen collection and processing procedures for these biomarkers followed the NHANES quality assurance and control (QA/QC) protocols that have been detailed elsewhere.[50-52] Briefly, the biomarker samples were obtained during the morning examination session after a 9-hour fast. After FPG was obtained, participants were asked to drink a calibrated dose of Trutol™ (~75 g glucose), and plasma glucose was again assessed two hours after the oral glucose load was administered.[53] Glucose levels were determined by a hexokinase method,[54] while HbA1c was measured by high-performance liquid chromatography.[55]

Risk-based Screening to Identify Individuals with Pre-diabetes and Diabetes

We applied both the ADA/CDC adult self-assessment screener and the AAP/ADA pediatric clinical screening guidelines using variables available in NHANES (**Tables 1** and **2**). We then evaluated their performance in determining pre-DM/DM status when compared to the determinations using the ADA diagnostic biomarkers (elevated FPG, 2hrPG, or HbA1c). Both the adult screener and the AAP/ADA guidelines include weight status, family history of diabetes, hypertension status, and history of gestational diabetes (history of maternal gestational diabetes for the youth screening guidelines). We calculated BMI percentiles using the SAS program provided by the CDC .[56] Since NHANES does not ask children directly about family history of type 2 diabetes in first- or second-degree relatives, we used the response to the question, "[Have you/Has SP] ever been told by a doctor or other health professional that [you have/s/he has] health conditions or a medical or family history that increases

medRxiv preprint doi: https://doi.org/10.1101/19007872.this version posted October 5, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

[your/his/her] risk for diabetes?", to assess family history of diabetes. We calculated blood pressure percentiles for children (<18 years) based on the 2017 Clinical Practice Guidelines from the AAP using the recommended published SAS program.[57,58] We defined hypertension using blood pressure cutoffs according to the 2017 guidelines (≥90th percentile or ≥120/80 mmHg for children <13 years, and ≥120/80 for those ≥13 years).[57] Information about personal/maternal history of gestational diabetes was not available for children, and was thus not included. The AAP/ADA pediatric clinical screening guidelines also include other symptoms or conditions related to insulin resistance, namely dyslipidemia, presence of acanthosis nigricans, prior diagnosis of polycystic ovary syndrome (PCOS), and small-for-gestational-age birth weight. We defined dyslipidemia as elevated total cholesterol level (≥170 mg/dL) according to AAP guidelines.[59] Although NHANES also included measurement of serum cholesterol low-density lipoprotein (LDL), we did not use it to define dyslipidemia due to substantial amounts of missing data. We were not able to find variables related to acanthosis nigricans, PCOS, or birth weight in NHANES data, and thus did not include these variables. The adult ADA/CDC screener also includes a question about physical activity. We defined a child as physically active if he/she had any of the following responses to questions from the physical activity questionnaire (PAQ) in NHANES: Times per week play or exercise hard (≥ 1 hour); Days physically active at least 60 min (≥1); Vigorous work activity (yes); Moderate work activity (yes); Days walk or bicycle (>1); Vigorous recreational activities (Yes); Moderate recreational activities (Yes); Muscle strengthening activities (Yes); Tasks around home/yard (Yes); average level of physical activity each day (>not much).

Assessment of the Performance of the Adult pre-DM/DM Screener and Pediatric Screening Guidelines

As the purpose of the current study was to evaluate the performance of the adult ADA/CDC screener and the AAP/ADA pediatric clinical screening guidelines in identifying pre-DM/DM, not to make population level estimates of disease prevalence, we did not apply survey procedures to the NHANES data and reported the unweighted results. Analyses were conducted in SAS (v9.4). We applied McNemar's test to assess the agreement between determinations of pre-DM/DM status made by the adult screener/pediatric screening guidelines and the ADA biomarker diagnostic criteria, and also reported Kappa coefficients (lower value indicates lower agreement). We conducted these assessments on the overall NHANES data set collected in our study, as well as sub-datasets stratified by sex (male, female), race/ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, other), and age groups (12-14 years, 15-17 years, and 18-19 years). We also tested whether the agreement between the pediatric screening guidelines and the biomarker diagnostic criteria (i.e. the marginal homogeneity) differed across subgroups using the Breslow-Day test and test for equal Kappa coefficients.

Machine Learning to Classify pre-DM/DM Status

As alternatives to expert-defined screeners, we also explored automated algorithms from machine learning [41] for developing pre-DM/DM classifiers directly from the youth NHANES data. The goal of these algorithms is to automatically infer a mathematical function that can categorize individuals, described in terms of relevant

medRxiv preprint doi: https://doi.org/10.1101/19007872.this version posted October 5, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

features, into classes, here pre-DM/DM or not. Due to the higher relevance of the AAP/ADA pediatric clinical screening guidelines to youth pre-DM/DM, we used the same five variables used in these guidelines, namely continuous BMI percentiles, family history of diabetes (yes/no), race ethnicity (non-Hispanic white vs otherwise), hypertension (yes/no), and continuous total cholesterol levels, as features for developing these classifiers. The algorithms we considered for this purpose were AdaBoost(M1), LogitBoost, Naïve Bayes, Logistic (Regression), Support Vector Machine (SMO), Voted Perceptron, K-nearest neighbor (IBk), PART and J48 decision tree inference algorithms, and Random Forest [41]. Their implementations in the Weka ML toolkit [60], incorporated into our DataSink pipeline [61], were used with their default parameter values in Weka. Since these implementations produce probabilistic predictions, we used the value of the probability score that yields the highest value of balanced accuracy, the mean of specificity and sensitivity [62], as the threshold to convert them into binary predictions to be comparable with the screening test results. All these ML algorithms were tested in a five-fold cross-validation (CV) setup [63], where in each CV round, 80% of the data were used to train candidate classifiers, which were then used to make pre-DM/DM predictions on the remaining 20% of the data. To address the imbalance between the pre-DM/DM and non-pre-DM/DM classes in the data (**Table 3**), which can adversely affect the classifiers [64], our framework balanced the two classes by randomly under-sampling the majority class during training [64], while retaining the distribution in the test fold exactly the same. Repeating this process over the whole dataset, and collecting the predictions for all five subsets of the data created during CV, generated dataset-wide predictions from all the algorithms, which were then evaluated as described below.

Evaluation of pre-DM/DM Status Determined by Clinical Screeners and Machine Learning-based Classifiers

Both the adult ADA/CDC screener and the pediatric clinical screening guidelines, as well as the ML algorithms described above, produce binary classifications, specifically positive (+) for pre-DM/DM determinations, and negative (-) for non-pre-DM/DM ones. Although several measures are known for evaluating such classifications against true labels [65], determined here based on ADA biomarker measurements, the most commonly used measures like accuracy and AUC score are not ideal for our data. For instance, since only about 30% of the youth in the NHANES population considered in our study are pre-DM/DM cases, one can get an artificially high accuracy value of about 70% just by classifying all the individuals as not having pre-DM/DM, which is obviously not useful. The AUC score, which is designed for continuous predictions, not binary ones, and thus isn't directly applicable in our case, is also not informative in cases of class imbalance like ours[66].

Based on these considerations, we focused on evaluation measures that have been designed for and shown to be the most relevant for classifier evaluation in the case of unbalanced classes [66]. Specifically, we used six measures [65] to evaluate the correctness of the pre-DM/DM predictions from the screening guidelines and ML methods: sensitivity (recall+), specificity (recall-), positive predictive value (PPV, precision+), negative predictive value (NPV, precision-), and F-measures for the two classes (see **Table 4** for detailed definitions of these measures). Another advantage of

medRxiv preprint doi: https://doi.org/10.1101/19007872.this version posted October 5, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

these measures, as compared to accuracy and AUC score, is that they provide per-class performance assessment, which can reveal if the classifier is biased in favor of a particular class, which is likely in the case of unbalanced classes[66].

Finally, we conducted the recommended Friedman and Nemenyi tests [67] to assess the statistical significance of the comparisons of the predictive performances of all the ML methods tested, as well as that of the pediatric screening guidelines.

Results

Study population: Descriptive demographic and clinical characteristics of the 2858 youth in the study population are presented in **Table 3**.

Comparisons of pre-DM/DM status obtained from adult screener/pediatric screening guidelines against pre-DM/DM determinations using biomarker diagnostic criteria: The number and percentages of youth in our study population found to have pre-DM/DM according to the ADA/CDC adult screener and the AAP/ADA pediatric clinical screening guidelines are shown in **Table 3**. We found poor agreement between the pre-DM/DM determinations based on biomarker criteria and the adult screener (McNemar's test p-value<0.0001, Kappa coefficient 0.07 (95%CI: 0.04-0.09)), as well as between those based on biomarker criteria and the pediatric clinical screening guidelines (McNemar's test p<0.0001, Kappa coefficient 0.1 (95%CI: 0.06-0.14)).

Predictive performance of the adult ADA/CDC screener and the AAP/ADA pediatric clinical screening guidelines are shown in **Table 4**. The top part of the table shows the contingency tables detailing these comparisons, while the bottom part lists the variety of performance measures derived from these tables. These measures were chosen because of their relevance to cases where the classes (pre-DM/DM or not) are unbalanced, and thus, more routinely used measures like accuracy and AUC may not be informative.[68] The adult screener performed poorly in identifying youth with pre-DM/DM (sensitivity=7.2%, specificity=97.7%; **Table 4**). The performance of the pediatric guidelines was better, as it correctly identified 43.1% of the youth with pre-DM/DM based on biomarkers; the PPV (precision+) was 35.2%, and the pre-DM/DM F-measure was 38.8% (**Table 4**).

Assessment of clinical screening guidelines' performance across demographic subgroups: The agreement between the pediatric clinical screening guidelines and the biomarker criteria in defining pre-DM/DM status differed significantly for males and females (Breslow-Day Test p=0.02), and had a borderline statistically significant difference across the three age groups (Breslow-Day Test p=0.046), but did not differ across the four racial/ethnic groups (Breslow-Day Test p=0.42). The Kappa coefficients did not differ significantly across all the subgroups (p>0.05).

Figure 1 shows variations in the predictive performance measures of the clinical screening guidelines across the various subgroups studied. The sensitivity (recall+) of the clinical screening guidelines in identifying pre-DM/DM was higher among females than males (52.2% vs 38.2%), while the PPV (precision+) was lower among females (29.4% vs 41.1%). The guidelines performed better for Hispanics and non-Hispanic Blacks than for non-Hispanic Whites and other racial/ethnic groups in terms of

medRxiv preprint doi: https://doi.org/10.1101/19007872.this version posted October 5, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

sensitivity (51.8% and 51.9% vs 23.4% and 32.5%), while the PPV was similar (28.8%-37.6%) across the four racial/ethnic groups. Finally, the guidelines performed the worst for those aged 12-14 years (sensitivity=39.9%) and the best for those aged 18-19 years (sensitivity=47.8%, PPV=30.2%, and F-measure=43.7%).

Machine learning to classify pre-diabetes: **Figure 2** shows the five-fold cross-validation-derived results of classifying pre-DM/DM status using ML algorithms, variables used in the pediatric screening guidelines, and class labels (pre-DM/DM or not) defined using the ADA biomarker diagnostic guidelines. It can be seen from these results that, across almost all the algorithms and evaluation measures, it is comparatively easier to produce more accurate predictions for the non-pre-DM/DM class than the pre-DM/DM one. This is primarily due to the smaller number of pre-DM/DM-diagnosed individuals that can contribute to the inference of the ML-based predictive models. Even so, the overall performance varies in a manner consistent with that of the clinical screening guidelines across the measures and classes. Furthermore, in each case, at least one ML algorithm performed statistically significantly better than the screening guidelines, especially for the harder to predict pre-DM/DM class. In particular, the Naïve Bayes classifier performed equivalently or significantly better as compared to the screening guidelines in terms of all the measures for this class (Friedman-Nemenyi test $p=9.216\times10^{-5}$, 0.252 and 5.228×10^{-5} for PPV, sensitivity and F-measure respectively). This classifier also performed significantly better than the screening guidelines for the non-pre-DM/DM class (Friedman-Nemenyi test $p=8.5\times10^{-10}$, 0.013 and 0.005 for NPV, specificity and F-measure respectively). Several other algorithms, such as Logistic (Regression), LogitBoost, PART and J48 (decision tree), also performed statistically equivalently or significantly better than the screening guidelines.

Overall, these results show that even with very few features (only five here), data-driven ML-based methods can help improve upon the performance of traditional pre-DM/DM screening guidelines.

Discussion

While there are simple, non-invasive, questionnaire-based patient self-assessment tools to identify adults at-risk for pre-DM/DM, no such tools exist for youth. The increasing incidence and prevalence of type 2 diabetes and pre-diabetes among youth in recent years, and the under-diagnosis of these conditions despite serious long-term sequelae, point to a pressing need for the development of simple accurate screening tools for identifying at-risk youth. Application of these tools can guide patients, families, and providers in conducting follow-up clinical tests for definitive diagnosis and targeting resource-intense diabetes prevention programs. Towards that end, we used NHANES data to evaluate the performance of an existing adult screening tool recommended by the ADA and CDC, as well as current pediatric clinical screening guidelines recommended by the AAP and ADA, against diagnosis of pre-DM/DM based on ADA biomarker diagnostic criteria (elevated FPG/2hrPG/HbA1C). In addition, given the differences in the prevalence of pre-DM/DM among youth based on race/ethnicity, age, and sex, we examined the performance of the clinical screening guidelines among youth from different subgroups based on these demographic characteristics. Overall,

medRxiv preprint doi: https://doi.org/10.1101/19007872.this version posted October 5, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

these two screeners did not perform well in identifying youth with pre-DM/DM, and there were notable differences in the performance of the clinical screening guidelines across the above demographic subgroups.

The poor performance of the adult ADA/CDC screener in identifying youth with pre-DM/DM is not unexpected, as the screener was developed for the adult population. The below 10% sensitivity of this screener found in our study for youth stood in contrast with a much higher sensitivity (about 80%) found in studies evaluating the screener on adult NHANES data.[34,69,70] Other adult screeners that include many of the same criteria as the ADA/CDC screener (and some additional criteria such as waist circumference and fruit/vegetable intake) have also been shown to perform well in adults based on evaluation of NHANES data.[71-73] However, these screeners were also not developed for youth, highlighting the need to develop a simple questionnaire-based screener specifically for the youth population.

Despite the fact that the AAP/ADA pediatric clinical screening guidelines are meant for health care providers to identify youth at risk for diabetes, the sensitivity of these guidelines in identifying NHANES youth with pre-DM/DM was still below 50%. In addition, as there are significant disparities in youth pre-DM/DM with higher rates among racial/ethnic minorities, males, and older adolescents, it is important to accurately identify risk in different demographic subgroups. Use of the unweighted NHANES data, although not nationally representative, offers an advantage here, since NHANES oversamples racial/ethnic minority populations which have higher rates of pre-DM and DM. We found that the clinical guidelines performed slightly better in females than males, and across age groups, while the performance didn't differ significantly across race/ethnicity subgroups. Other studies have examined the performance of risk tools across sex, race/ethnicity, and age, and have found differences in performance across the corresponding subgroups.[36,74,75] One study that examined the performance of an adult diabetes risk tool across sex and race/ethnicity found that the tool performed better in women than in men and in non-Hispanic Whites than in Hispanics in detecting undiagnosed pre-DM/DM.[36] This points to the need to evaluate the performance of risk tools in different demographic subgroups and perhaps develop different subgroup-specific tools.

Data-driven ML-based algorithms yielded slight but statistically significant improvements in identifying youth with pre-DM/DM, despite using only five variables (BMI, family history of diabetes, race/ethnicity, hypertension, and cholesterol levels). Combining many more relevant features from NHANES or other large data sets with rich clinical and behavioral health data, as well as powerful ML approaches like feature selection[76] and deep learning[77], is likely to substantially enhance our ability to develop a data-driven, relatively simple, and accurate screener for youth at risk for pre-DM/DM.

Our study has several notable limitations. We defined the diagnosis of pre-DM/DM based on one-time measurements of biomarkers, whereas the ADA recommends repeated measurements.[12] In addition, our application of the adult ADA/CDC screener to the youth NHANES data was not perfect, as there were some variables that were

medRxiv preprint doi: https://doi.org/10.1101/19007872.this version posted October 5, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

either not available in NHANES (history of gestational DM) or available in a slightly different format (family history, blood pressure, weight status, and physical activity). Interpretation of age (all youth were assigned a score of zero for age because the screener uses adult age group cutoffs (as shown in Table 1) and sex (a score of 1 for males and 0 for females) may not reflect true age and sex differences affecting youth pre-DM/DM risk. Similarly, we were not able to exactly apply the AAP/ADA pediatric clinical screening guidelines because of missing information (history of maternal gestational diabetes during the child's gestation, presence of acanthosis nigricans, diagnosis of polycystic ovary syndrome, and history of small-for-gestational-age birthweight), or information available in a different format (family history of diabetes). Also, because the adult screener and the pediatric clinical guidelines are only meant to identify increased risk of diabetes and not to distinguish risk of pre-DM from risk of DM, we did not explicitly differentiate between these conditions in our study. Specifically, we used biomarker cutoffs for pre-DM and examined the ability of existing screener/clinical guidelines to identify the presence of pre-DM/DM. The benefit of identifying pre-DM in particular is that, unlike DM, pre-DM is thought to be reversible with lifestyle modification and weight management.[4,5] Thus, accurate identification of pre-DM specifically may help better target intensive diabetes prevention programs. Furthermore, NHANES data, and thus, our evaluation did not differentiate type 1 vs type 2 diabetes, although the prevalence of type 1 diabetes is relatively low among youth compared to that of pre-DM and type 2 DM.[7,8] Finally, we only applied a routinely used set of classification algorithms that have been developed in machine learning[40,41], and our results could have been improved further by employing other more advanced algorithms.

Despite these limitations, our study also has several strengths. To our knowledge, this is the first study that examined the performance of a widely used adult diabetes risk screener, as well as recommended pediatric clinical screening guidelines, in identifying pre-DM/DM status, determined using biomarker diagnostic criteria, among youth. Our demonstration that neither of these tools performed well for this task points to the need for additional work to develop a simple yet accurate screener for youth diabetes risk. Studies to date that have focused on assessing youth pre-DM/DM risk have relied on small sample sizes from localized clinical settings, and have sometimes included invasive blood tests that may not be the best initial strategy to identify risk.[78,79] In contrast, NHANES includes a large sample of individuals from across the United States, including well-represented age, sex, and racial/ethnic subgroups, as well as detailed biomarker, clinical, and behavioral health data. While NHANES data have been used to develop diabetes risk screeners for adults,[34,70,71] and to examine prevalence of pre-DM/DM among youth,[8,80] no studies to date have used these data to develop and evaluate youth diabetes risk screeners. In particular, our investigation of machine learning methods applied to these data demonstrates the promise of automated data-driven methods for developing such screeners. Future work includes the use of more advanced ML methods to a wider range of clinical and behavioral health data available in NHANES to build better predictive tools for discriminating pre-DM/DM status. Such tools may be used by youth or their caretakers (parents, teachers, etc.), as well as in clinical and community settings to identify at-risk youth who can benefit from more intensive diabetes prevention programs.

medRxiv preprint doi: https://doi.org/10.1101/19007872.this version posted October 5, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

REFERENCES

1. Lotfy M, Adeghate J, Kalasz H, Singh J, Adeghate E. Chronic Complications of Diabetes Mellitus: A Mini Review. *Curr Diabetes Rev.* 2017;13(1):3-10.
2. Economic Costs of Diabetes in the U.S. in 2017. *Diabetes Care.* 2018;41(5):917-928.
3. Perreault L, Faerch K. Approaching pre-diabetes. *J Diabetes Complications.* 2014;28(2):226-233.
4. Weiss R, Taksali SE, Tamborlane WV, Burgert TS, Savoye M, Caprio S. Predictors of changes in glucose tolerance status in obese youth. *Diabetes care.* 2005;28(4):902-909.
5. Love-Osborne KA, Sheeder JL, Nadeau KJ, Zeitler P. Longitudinal follow up of dysglycemia in overweight and obese pediatric patients. *Pediatr Diabetes.* 2018;19(2):199-204.
6. Mayer-Davis EJ, Lawrence JM, Dabelea D, et al. Incidence Trends of Type 1 and Type 2 Diabetes among Youths, 2002-2012. *N Engl J Med.* 2017;376(15):1419-1429.
7. Dabelea D, Mayer-Davis EJ, Saydah S, et al. Prevalence of type 1 and type 2 diabetes among children and adolescents from 2001 to 2009. *Jama.* 2014;311(17):1778-1786.
8. Menke A, Casagrande S, Cowie CC. Prevalence of Diabetes in Adolescents Aged 12 to 19 Years in the United States, 2005-2014. *JAMA.* 2016;316(3):344-345.
9. Group TS, Zeitler P, Hirst K, et al. A clinical trial to maintain glycemic control in youth with type 2 diabetes. *The New England journal of medicine.* 2012;366(24):2247-2256.
10. Dart AB, Martens PJ, Rigatto C, Brownell MD, Dean HJ, Sellers EA. Earlier onset of complications in youth with type 2 diabetes. *Diabetes care.* 2014;37(2):436-443.
11. Nadeau KJ, Anderson BJ, Berg EG, et al. Youth-Onset Type 2 Diabetes Consensus Report: Current Status, Challenges, and Priorities. *Diabetes Care.* 2016;39(9):1635-1642.
12. Arslanian S, Bacha F, Grey M, Marcus MD, White NH, Zeitler P. Evaluation and Management of Youth-Onset Type 2 Diabetes: A Position Statement by the American Diabetes Association. *Diabetes Care.* 2018;41(12):2648-2668.
13. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2019. *Diabetes Care.* 2019;42(Suppl 1):S13-s28.
14. Bloomgarden ZT. Type 2 diabetes in the young: the evolving epidemic. *Diabetes Care.* 2004;27(4):998-1010.
15. Lee AM, Fermin CR, Filipp SL, Gurka MJ, DeBoer MD. Examining trends in prediabetes and its relationship with the metabolic syndrome in US adolescents, 1999-2014. *Acta Diabetol.* 2017;54(4):373-381.
16. Black LI, Nugent CN, Vahratian A. Access and Utilization of Selected Preventive Health Services Among Adolescents Aged 10-17. *NCHS Data Brief.* 2016(246):1-8.
17. Harris SK, Aalsma MC, Weitzman ER, et al. Research on Clinical Preventive Services for Adolescents and Young Adults: Where Are We and Where Do We Need to Go? *J Adolesc Health.* 2017;60(3):249-260.
18. Rhodes ET, Finkelstein JA, Marshall R, Allen C, Gillman MW, Ludwig DS. Screening for type 2 diabetes mellitus in children and adolescents: attitudes, barriers, and practices among pediatric clinicians. *Ambul Pediatr.* 2006;6(2):110-114.
19. Anand SG, Mehta SD, Adams WG. Diabetes mellitus screening in pediatric primary care. *Pediatrics.* 2006;118(5):1888-1895.
20. Lee JM, Eason A, Nelson C, Kazzi NG, Cowan AE, Tarini BA. Screening practices for identifying type 2 diabetes in adolescents. *J Adolesc Health.* 2014;54(2):139-143.
21. Lee JM, Gebremariam A, Wu EL, LaRose J, Gurney JG. Evaluation of nonfasting tests to screen for childhood and adolescent dysglycemia. *Diabetes care.* 2011;34(12):2597-2602.
22. Brackney DE, Cutshall M. Prevention of type 2 diabetes among youth: a systematic review, implications for the school nurse. *J Sch Nurs.* 2015;31(1):6-21.

medRxiv preprint doi: https://doi.org/10.1101/19007872.this version posted October 5, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

23. McGavock J, Dart A, Wicklow B. Lifestyle therapy for the treatment of youth with type 2 diabetes. *Curr Diab Rep.* 2015;15(1):568.
24. McCurley JL, Crawford MA, Gallo LC. Prevention of Type 2 Diabetes in U.S. Hispanic Youth: A Systematic Review of Lifestyle Interventions. *Am J Prev Med.* 2017;53(4):519-532.
25. Knowler WC, Fowler SE, Hamman RF, et al. 10-year follow-up of diabetes incidence and weight loss in the Diabetes Prevention Program Outcomes Study. *Lancet.* 2009;374(9702):1677-1686.
26. Brown N, Critchley J, Bogowicz P, Mayige M, Unwin N. Risk scores based on self-reported or available clinical data to detect undiagnosed type 2 diabetes: a systematic review. *Diabetes Res Clin Pract.* 2012;98(3):369-385.
27. Collins GS, Mallett S, Omar O, Yu LM. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. *BMC Med.* 2011;9:103.
28. Noble D, Mathur R, Dent T, Meads C, Greenhalgh T. Risk models and scores for type 2 diabetes: systematic review. *Bmj.* 2011;343:d7163.
29. Schwarz PE, Li J, Lindstrom J, Tuomilehto J. Tools for predicting the risk of type 2 diabetes in daily practice. *Horm Metab Res.* 2009;41(2):86-97.
30. Barber SR, Davies MJ, Khunti K, Gray LJ. Risk assessment tools for detecting those with pre-diabetes: a systematic review. *Diabetes Res Clin Pract.* 2014;105(1):1-13.
31. Abbasi A, Peelen LM, Corpeleijn E, et al. Prediction models for risk of developing type 2 diabetes: systematic literature search and independent external validation study. *Bmj.* 2012;345:e5900.
32. Thoopputra T, Newby D, Schneider J, Li SC. Survey of diabetes risk assessment tools: concepts, structure and performance. *Diabetes Metab Res Rev.* 2012;28(6):485-498.
33. Buijsse B, Simmons RK, Griffin SJ, Schulze MB. Risk assessment tools for identifying individuals at risk of developing type 2 diabetes. *Epidemiol Rev.* 2011;33:46-62.
34. Bang H, Edwards AM, Bomback AS, et al. Development and validation of a patient self-assessment score for diabetes risk. *Ann Intern Med.* 2009;151(11):775-783.
35. Prediabetes Risk Test. American Diabetes Association and Centers for Disease Control and Prevention. https://www.cdc.gov/diabetes/prevention/pdf/Prediabetes-Risk-Test-Final.pdf. Accessed September 26, 2019.
36. Zhang L, Zhang Z, Zhang Y, Hu G, Chen L. Evaluation of Finnish Diabetes Risk Score in screening undiagnosed diabetes and prediabetes among U.S. adults by gender and race: NHANES 1999-2010. *PLoS One.* 2014;9(5):e97865.
37. Zipf G, Chiappa M, Porter KS, Ostchega Y, Lewis BG, Dostal J. National health and nutrition examination survey: plan and operations, 1999-2010. *Vital Health Stat 1.* 2013(56):1-37.
38. Menke A, Casagrande S, Geiss L, Cowie CC. Prevalence of and Trends in Diabetes Among Adults in the United States, 1988-2012. *Jama.* 2015;314(10):1021-1029.
39. Li C, Ford ES, Zhao G, Mokdad AH. Prevalence of pre-diabetes and its association with clustering of cardiometabolic risk factors and hyperinsulinemia among U.S. adolescents: National Health and Nutrition Examination Survey 2005-2006. *Diabetes care.* 2009;32(2):342-347.
40. Cleophas TJ, Zwinderman AH. *Machine Learning in Medicine-a Complete Overview.* Springer; 2016.
41. Alpaydin E. *Introduction to machine learning.* MIT press; 2014.
42. Deo RC. Machine Learning in Medicine. *Circulation.* 2015;132(20):1920-1930.
43. Gao M, Igata H, Takeuchi A, Sato K, Ikegaya Y. Machine learning-based prediction of adverse drug effects: An example of seizure-inducing compounds. *J Pharmacol Sci.* 2017;133(2):70-78.
44. Leung MKK, Delong A, Alipanahi B, Frey BJ. Machine Learning in Genomic Medicine: A Review of Computational Problems and Data Sets. *Proceedings of the IEEE.* 2016;104(1):176-197.
45. Steyerberg EW. *Clinical prediction models: a practical approach to development, validation, and updating.* Springer Science & Business Media; 2008.

medRxiv preprint doi: https://doi.org/10.1101/19007872.this version posted October 5, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

46. Pandey G, Pandey OP, Rogers AJ, et al. A Nasal Brush-based Classifier of Asthma Identified by Machine Learning Analysis of Nasal RNA Sequence Data. *Scientific Reports.* 2018;8(1):8826.
47. Varghese B, Chen F, Hwang D, et al. Objective risk stratification of prostate cancer using machine learning and radiomics applied to multiparametric magnetic resonance images. *Scientific Reports.* 2019;9(1):1570.
48. National Center for Health Statistics. How has the National Health and Nutrition Examination Survey (NHANES) improved the health of all Americans? https://www.cdc.gov/nchs/nhanes/dataaccomp.htm. Published 2015. Accessed July 23, 2019.
49. National Center for Health Statistics. NHANES Questionnaires, Datasets, and Related Documentation. https://wwwn.cdc.gov/nchs/nhanes/default.aspx. Published 2018. Accessed July 23, 2019.
50. National Center for Health Statistics. NHANES Survey Methods and Analytic Guidelines. https://wwwn.cdc.gov/nchs/nhanes/analyticguidelines.aspx#plan-and-operations. Published 2018. Accessed July 23, 2019.
51. National Center for Health Statistics. National Health and Nutrition Examination Survey (NHANES) MEC Laboratory Procedures Manual. https://wwwn.cdc.gov/nchs/data/nhanes/2013-2014/manuals/2013_mec_laboratory_procedures_manual.pdf. Published 2013. Accessed July 23, 2019.
52. National Center for Health Statistics. NHANES 2015-2016 Laboratory Data Overview. https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/overviewlab.aspx?BeginYear=2015. Published 2018. Accessed July 23, 2019.
53. National Center for Health Statistics. 2005-2006 Data Documentation, Codebook, and Frequencies Oral Glucose Tolerance Test (OGTT_D). https://wwwn.cdc.gov/Nchs/Nhanes/2005-2006/OGTT_D.htm. Published 2008. Accessed July 23, 2019.
54. Health NCf, Statistics. Laboratory Procedure Manual Fasting Glucose in Plasma

NHANES 2015-2016. https://wwwn.cdc.gov/nchs/data/nhanes/2015-2016/labmethods/GLU_I_MET.pdf. Accessed.

55. National Center for Health Statistics. Glycohemoglobin in Whole Blood using Tosoh G 7 HPLC. https://www.cdc.gov/nchs/data/nhanes/nhanes_07_08/GHB_E_met_Tosoh_G7.pdf. Published 2008. Accessed July 23, 2019.
56. Centers for Disease Control and Prevention. A SAS Program for the 2000 CDC Growth Charts (ages 0 to < 20 years). https://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm. Published 2019. Accessed July 23, 2019.
57. Flynn JT, Kaelber DC, Baker-Smith CM, et al. Clinical Practice Guideline for Screening and Management of High Blood Pressure in Children and Adolescents. *Pediatrics.* 2017;140(3).
58. BernardRosner. Childhood Blood Pressure Macro-batch mode. https://sites.google.com/a/channing.harvard.edu/bernardrosner/pediatric-blood-press/childhood-blood-pressure. Published 2018. Accessed July 23, 2019.
59. American Academy of Pediatrics. Cholesterol Levels in Children and Adolescents. https://www.healthychildren.org/English/healthy-living/nutrition/Pages/Cholesterol-Levels-in-Children-and-Adolescents.aspx. Published 2019. Accessed July 23, 2019.
60. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA data mining software: an update. *SIGKDD Explor Newsl.* 2009;11(1):10-18.
61. Whalen S, Pandey OP, Pandey G. Predicting protein function and other biomedical characteristics with heterogeneous ensembles. *Methods.* 2016;93:92-102.
62. Sokolova M, Japkowicz N, Szpakowicz S. Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. Paper presented at: Australasian joint conference on artificial intelligence2006.
63. Arlot S, Celisse A. A survey of cross-validation procedures for model selection. *Statist Surv.* 2010;4:40-79.

medRxiv preprint doi: https://doi.org/10.1101/19007872.this version posted October 5, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

64. Chawla NV. Data Mining for Imbalanced Datasets: An Overview. In: Maimon O, Rokach L, eds. *Data Mining and Knowledge Discovery Handbook.* Boston, MA: Springer US; 2005:853-867.
65. Lever J, Krzywinski M, Altman N. Points of Significance: Classification evaluation. *Nat Meth.* 2016;13(8):603-604.
66. Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One.* 2015;10(3):e0118432.
67. Demsar J. Statistical Comparisons of Classifiers over Multiple Data Sets. *J Mach Learn Res.* 2006;7:1-30.
68. !!! INVALID CITATION !!! (65, 66).
69. Poltavskiy E, Kim DJ, Bang H. Comparison of screening scores for diabetes and prediabetes. *Diabetes Res Clin Pract.* 2016;118:146-153.
70. Herman WH, Smith PJ, Thompson TJ, Engelgau MM, Aubert RE. A new and simple questionnaire to identify people at increased risk for undiagnosed diabetes. *Diabetes Care.* 1995;18(3):382-387.
71. Heikes KE, Eddy DM, Arondekar B, Schlessinger L. Diabetes Risk Calculator: a simple tool for detecting undiagnosed diabetes and pre-diabetes. *Diabetes Care.* 2008;31(5):1040-1045.
72. Yu W, Liu T, Valdez R, Gwinn M, Khoury MJ. Application of support vector machine modeling for prediction of common diseases: the case of diabetes and pre-diabetes. *BMC Med Inform Decis Mak.* 2010;10:16.
73. Zhang Y, Hu G, Zhang L, Mayo R, Chen L. A novel testing model for opportunistic screening of pre-diabetes and diabetes among U.S. adults. *PLoS One.* 2015;10(3):e0120382.
74. Gooding HC, Ning H, Gillman MW, et al. Application of a Lifestyle-Based Tool to Estimate Premature Cardiovascular Disease Events in Young Adults: The Coronary Artery Risk Development in Young Adults (CARDIA) Study. *JAMA Intern Med.* 2017;177(9):1354-1360.
75. Stockdale KC, Olver ME, Wong SC. The Psychopathy Checklist: Youth Version and adolescent and adult recidivism: considerations with respect to gender, ethnicity, and age. *Psychol Assess.* 2010;22(4):768-781.
76. Saeys Y, Inza I, Larranaga P. A review of feature selection techniques in bioinformatics. *Bioinformatics.* 2007;23(19):2507-2517.
77. Ching T, Himmelstein DS, Beaulieu-Jones BK, et al. Opportunities and obstacles for deep learning in biology and medicine. *Journal of The Royal Society Interface.* 2018;15(141):20170387.
78. Lee JM, Gebremariam A, Woolford SJ, et al. A risk score for identifying overweight adolescents with dysglycemia in primary care settings. *J Pediatr Endocrinol Metab.* 2013;26(5-6):477-488.
79. Santoro N, Amato A, Grandone A, et al. Predicting metabolic syndrome in obese children and adolescents: look, measure and ask. *Obes Facts.* 2013;6(1):48-56.
80. May AL, Kuklina EV, Yoon PW. Prevalence of cardiovascular disease risk factors among US adolescents, 1999-2008. *Pediatrics.* 2012;129(6):1035-1041.

TABLES

Table 1. Questions in the ADA/CDC adult screener used to define pre-DM/DM status and corresponding operationally defined equivalent variables in NHANES used in this study.

Adult ADA/CDC pre-DM/DM risk test (at-risk if summed score≥5)	**NHANES variables used** (at-risk if summed score≥5)
1. How old are you? ≤40 years: 0 points 40-49 years : 1 point 50-59 years: 2 points ≥60 years: 3 points	Age<20 years: 0 points
2. Are you a woman or man? Man: 1 point Woman: 0 points	Male: 1 point Female: 0 points
3. If you are a woman, have you ever been diagn with gestational diabetes? Yes: 1 point No: 0 points	Not available. 0 points
4. Do you have a mother, father, sister, or brother with diabetes? Yes: 1 point No: 0 points	Ever been told by a doctor or other health professional that you have health conditions or a medical or family history that increases your risk for diabetes? Yes: 1 point No: 0 points
5. Have you ever been diagnosed with high blood pressure? Yes: 1 point No: 0 points	Based on blood pressure (BP) or BP percentile[†]: BP≥90th percentile or BP≥120/80 mm Hg for children ≥13 years: 1 point BP<90th percentile or BP<120/80 mm Hg for children ≥13 years: 0 points
6. Are you physically active Yes: 0 points No: 1 point	Derived responses indicating child is physically active[‡] Yes: 0 points No: 1 point
7. What is your weight category? Not overweight or obese: 0 points Overweight: 1 point Obese: 2 points Extremely obese: 3 points	Based on BMI percentile[§]: BMI<85th percentile: 0 points 85th≤BMI≤ 95th percentile: 1 point 95th≤BMI≤ 99th percentile: 2 point BMI≥99th percentile: 3 points

medRxiv preprint doi: https://doi.org/10.1101/19007872.this version posted October 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Notes: †, We calculated blood pressure percentiles for children (<18 years) based on the 2017 Clinical Practice Guidelines from the AAP using the recommended published SAS program.[57,58]
‡, We defined a child as physically active if he/she had any of the following responses to questions from the physical activity questionnaire (PAQ) in NHANES: Times per week play or exercise hard (≥ 1 hour); Days physically active at least 60 min (≥1); Vigorous work activity (yes); Moderate work activity (yes); Days walk or bicycle (>1); Vigorous recreational activities (Yes); Moderate recreational activities (Yes); Muscle strengthening activities (Yes); Tasks around home/yard (Yes); average level of physical activity each day (>not much).
§, We calculated BMI percentiles using the SAS program provided by the CDC for the 2000 CDC growth charts (ages 0 to <20 years) with overweight/obese defined as ≥85th percentile.[56]

Table 2. AAP/ADA pediatric clinical screening guidelines used to define pre-DM/DM status and their corresponding operationally defined equivalent variables in NHANES used in this study.

ADA/AAP pre-DM/DM risk for children (at-risk if overweight plus one or more additional risk factors)	**NHANES variables used** (at-risk if overweight plus one or more additional risk factors)
Overweight (BMI >85th percentile for age and sex, weight for height >85th percentile, or weight >120% of ideal for height) **A**†	BMI ≥ 85th percentile‡
Additional risk factors:	Additional risk factors:
Maternal history of gestational diabetes during the child's gestation **A**†	Not available.
Family history of type 2 diabetes in first- or second-degree relative **A**†	Ever been told by a doctor or other health professional that you have health conditions or a medical or family history that increases your risk for diabetes?
Race/Ethnicity (Native American, African American, Latino, Asian American, Pacific Islander) **A**†	Non-White race/ethnicity (non-Hispanic Black, Hispanic, other)
Signs of insulin resistance or conditions associated with insulin resistance (hypertension, dyslipidemia, acanthosis nigricans, polycystic ovary syndrome, or small-for-gestational-age birth weight). **B**†	Hypertension§: BP≥90th percentile or BP≥120/80 mm Hg for children ≥13 years; Dyslipidemia¶: total cholesterol ≥170 mg/dL

Notes: †, Evidence grades, with grade A and B representing higher and moderate quality evidence, respectively. Grades do not factor into the determination of risk in the current study.
‡, We calculated BMI percentiles using the SAS program provided by the CDC for the 2000 CDC growth charts (ages 0 to <20 years) with overweight/obese defined as ≥85th percentile.[56]

medRxiv preprint doi: https://doi.org/10.1101/19007872 this version posted October 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[§], We calculated blood pressure percentiles for children (<18 years) based on the 2017 Clinical Practice Guidelines from the AAP using the recommended published SAS program.[57,58]
[¶], We defined dyslipidemia as elevated total cholesterol level (≥170 mg/dL) according to AAP guidelines.[59]

Table 3. Characteristics of the study population (n=2858).

	Unweighted frequency	**Unweighted Proportion (%)**
Sex		
Male	1505	52.66
Female	1353	47.34
Race/Ethnicity		
Non-Hispanic white	788	27.57
Non-Hispanic black	764	26.73
Hispanic	1041	36.42
other	265	9.27
Age group		
12-14 years	1071	37.47
15-17 years	1088	38.07
18-19 years	699	24.46
Diagnosed with pre-DM/DM based on elevated FPG/2hrPG/HbA1c		
No	2035	71.20
Yes	823	28.80
At risk for pre-DM/DM based on ADA/CDC adult screener		
No	2751	96.29
Yes	106	3.71
At risk for pre-DM/DM based on AAP/ADA pediatric clinical screening guidelines		
No	1844	64.52
Yes	1014	35.48
	Mean (SD)	**Median**
Age (years)	15.5 (2.3)	16.0

medRxiv preprint doi: https://doi.org/10.1101/19007872 this version posted October 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Fasting plasma glucose (FPG, mg/dL)	94.3 (8.1)	94.0
Two hour plasma glucose (2hrPG, mg/dL)	98.2 (23.5)	96.0
Glycohemoglobin (HbA1c, %)	5.2 (0.3)	5.2
Cholesterol, total (mg/dL)	159.3 (30.3)	156.0
BMI-for-age percentile	67.9 (28.9)	76.6

Table 4. Performance measures of adult ADA/CDC screener and AAP/ADA pediatric clinical screening guidelines when compared against pre-DM/DM determinations based on ADA biomarker diagnostic criteria.

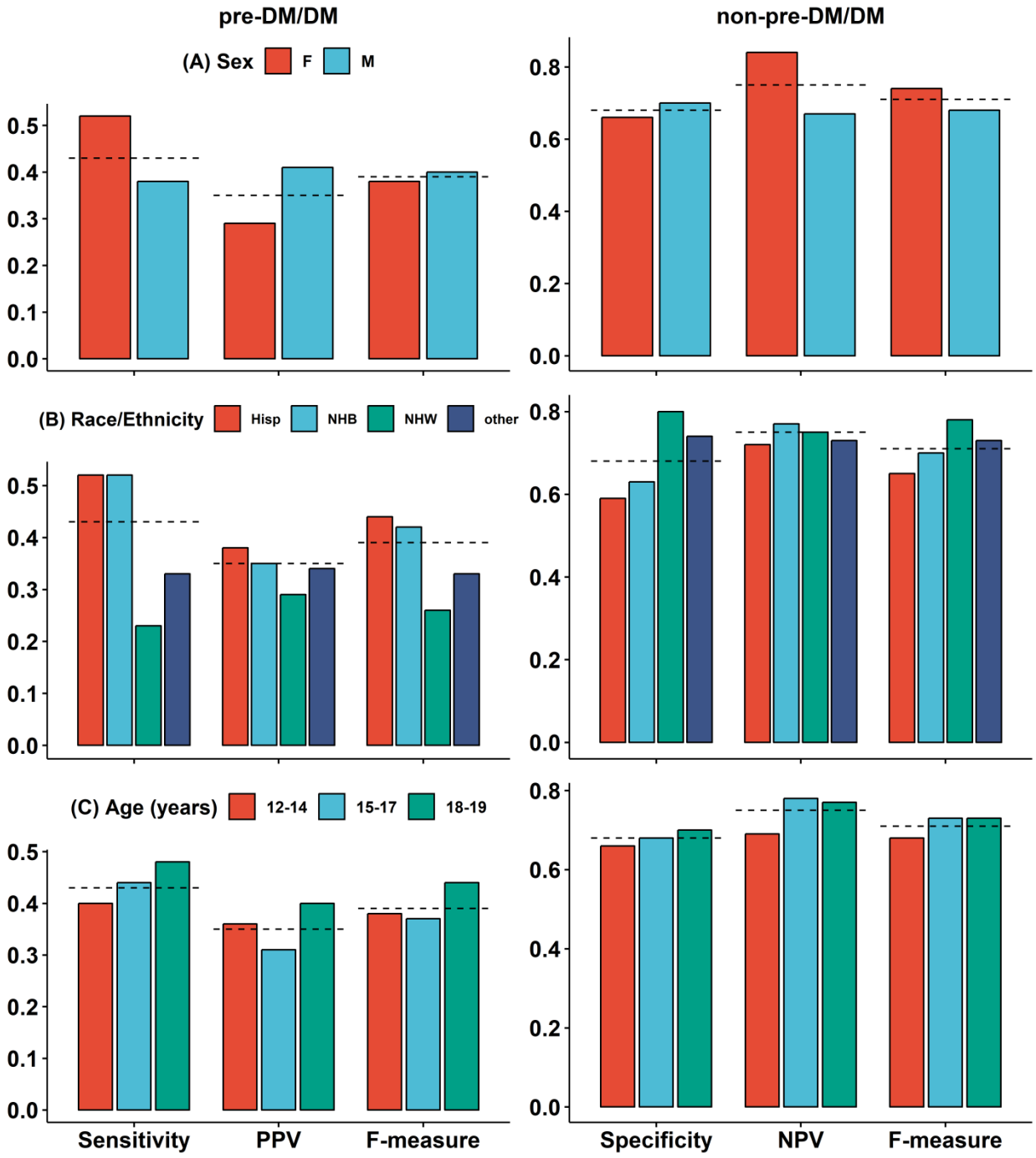
		Adult ADA/CDC screener for diabetes risk			**AAP/ADA pediatric clinical screening guidelines**		
Diagnosis of pre-DM/DM based on elevated FPG/2hrPG/HbA1C		**Yes**	**No**	*Row total*	**Yes**	**No**	*Row total*
	Yes	**60**	**768**	*828*	**357**	**471**	*828*
	No	**46**	**1984**	*2030*	**657**	**1373**	*2030*
	Column total	*106*	*2752*	*2858*	*1014*	*1844*	*2858*
Performance measures of the adult screener/pediatric clinical screening guidelines when compared against diagnosis of pre-DM/DM based on biomarkers for the positive (+) and negative class (-)							
Sensitivity (recall+) = Proportion at-risk based on adult screener/ pediatric clinical screening guidelines that have pre-DM/DM based on biomarkers		60/828 = 7.2%			357/828 = 43.1%		
Specificity (recall-) = Proportion not at-risk based on adult screener/ pediatric clinical screening guidelines that do not have pre-DM/DM based on biomarkers		1984/2030 =97.7%			1272/2030 = 67.6%		
Positive Predictive Value (PPV, precision+) = Proportion of truly at-risk (diagnosed with pre-DM/DM based on biomarkers) among all predicted to be at-risk based on adult screener/ pediatric clinical screening guidelines		60/106 = 56.6%			357/1014 = 35.2%		
Negative Predictive Value (NPV, precision-) = Proportion of truly not at-risk (not diagnosed with pre-DM/DM based on biomarkers) among all predicted not to be at-risk based on adult screener/ pediatric clinical screening guidelines		1984/2752 = 72.1%			1373/1844 = 74.5%		
F-measure+ = Harmonic (conservative) mean of Precision+ and Recall+ = 2*(Precision+ ***** Recall+) / (Precision+ **+** Recall+)		2*(7.2%*56.6%)/(7.2%+56.6%) = 12.8%			2*(43.1%*35.2%)/(43.1%+35.2%) = 38.8%		
F-measure- = Harmonic (conservative) mean of Precision- and Recall- = 2*(Precision- ***** Recall-) / (Precision- **+** Recall-)		2*(97.7%*72.1%)/(97.7%+72.1%) = 83.0%			2*(67.6%*74.5%)/(67.6%+74.5%) = 70.9%		

medRxiv preprint doi: https://doi.org/10.1101/19007872 this version posted October 5, 2019. The copyright holder for this preprint (**which was not certified by peer review**) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

medRxiv preprint doi: https://doi.org/10.1101/19007872.this version posted October 5, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

FIGURES

Figure 1. Variations in the performance of the American Diabetes Association pediatric screening guidelines in identifying youth with pre-diabetes/diabetes based on diagnostic biomarker measurements across subgroups based on (**A**) sex, (**B**) race/ethnicity and (**C**) age. Dashed lines denote the value of the corresponding evaluation measure obtained from the full study population (youth ages 12-19, National Health and Nutrition Examination Survey data, 2005-2016).



medRxiv preprint doi: https://doi.org/10.1101/19007872.this version posted October 5, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

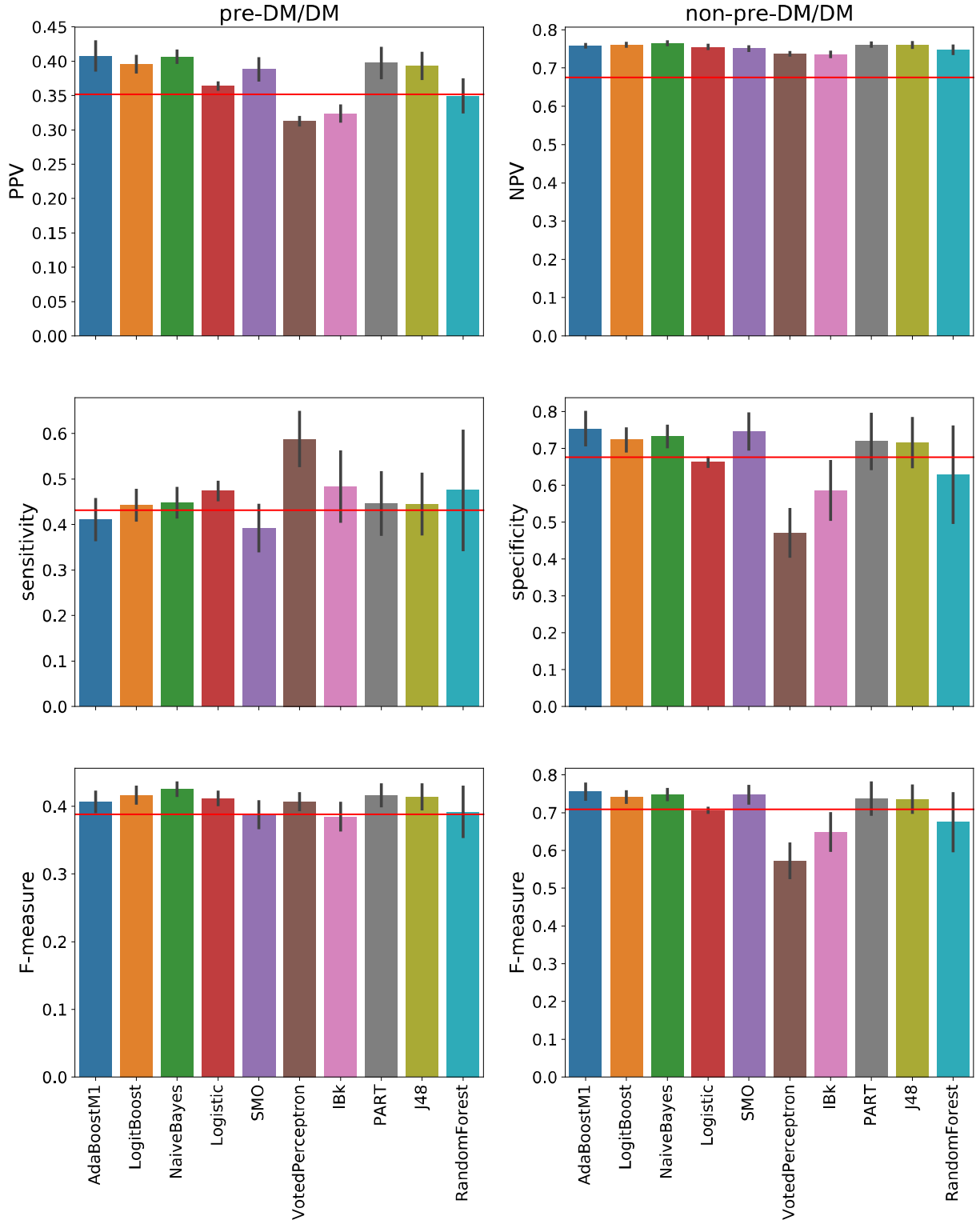



Figure 2. Performance of machine learning algorithms in classifying individuals into pre-DM/DM and non- pre-DM/DM classes, evaluated in terms of predictive value, sensitivity/specificity and F-measures for both classes. The variables used in this classification were the same as those used in the American Diabetes Association pediatric screening guidelines, whose performance in terms of each measure is shown by a horizontal red line in the corresponding subplot.