January 2020

External Validation Of An Electronic Phenotyping Algorithm To Detect Attention To Elevated Bmi And Weight-Related Comorbidities In Pediatric Primary Care.

Anya Golkowski Barron

External Validation of an Electronic Phenotyping Algorithm to Detect Attention to Elevated BMI and Weight-Related Comorbidities in Pediatric Primary Care

A Thesis Submitted to the
Yale University School of Medicine
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Medicine

By
Anya Golkowski Barron
2020

ABSTRACT

External Validation of an Electronic Phenotyping Algorithm to Detect Attention to Elevated BMI and Weight-Related Comorbidities in Pediatric Primary Care.

Anya Golkowski Barron[1], Christy Turer [2], Ada Fenick[1], Kaitlin Maciejewski[1], and Mona Sharifi[1].

[1]Department of Pediatrics, Yale University, School of Medicine, New Haven, CT.
[2]Department of Pediatrics, University of Texas Southwestern Medical Center and Children's Health, Dallas, TX.

Pediatric obesity is a growing national and global concern with nearly 1 in 5 children in the U.S. affected [1].The American Academy of Pediatrics endorsed expert committee recommendations in 2007 to assist clinicians in pediatric weight management; however, adherence to these recommendations among primary care providers is suboptimal, and measuring adherence in feasible and pragmatic ways is challenging[2-4]. Commonly used quality measures that rely on billing data alone are an inadequate measure of provider attention to weight status in pediatric populations as they do not capture whether providers communicate about elevated body mass index (BMI) and associated medical risks with families. Electronic phenotyping is a unique tool that has the ability to use multiple areas of stored clinical data to group individuals according to pre-defined characteristics such as diagnostic codes, laboratory values or medications. We examined the external validity of a phenotyping algorithm, developed previously by Turer et al and validated in a single health system in Texas, that assesses pediatric providers' attention to obesity and overweight using structured data from the electronic health record (EHR), to three pediatric primary care practices affiliated with Yale New Haven Health. Well child visit encounters were labeled either "no attention", "attention to BMI only", "attention to comorbidity only," or "attention to BMI and comorbidity". The performance of the algorithm was evaluated on the ability to predict "no attention", using

chart review as the reference standard. The application of the minimally altered algorithm yielded a sensitivity of 94.0% and a specificity of 79.2% for predicting “no attention”, compared to a sensitivity of 97.9% and a specificity of 94.8% in the original study. Our findings suggest that while electronic phenotyping using structured EHR inputs provides a better evaluation of clinic encounters than use of diagnostic codes alone, methods that incorporate information in unstructured (“free text”) clinical notes may yield better results.

Acknowledgements

To Julian and Brian for being the salt and light of my world.

To my parents for being the giants whose shoulders I stand on.

To Dr. Mona Sharifi the office of student research and the department of Pediatrics without whom this work would not be possible.

TABLE OF CONTENTS

INTRODUCTION

1. Definitions of Pediatric Overweight and Obesity

Overweight and obesity are clinical terms used to denote excess body weight, most frequently thought of in the form of adipose tissue. A commonly used measure for estimating body fat percentages in medicine is body mass index (BMI). BMI provides a measure of body weight adjusted for height, and although it does not provide a direct measure of body fat, levels do correlate with and are predictive of future adiposity [5]. BMI is also clinically useful as it can easily be assessed in the primary care setting with routine measurements of height and weight as opposed to more precise but less feasible methods such as dual-energy x-ray absorptiometry. Given the nature of the calculation, BMI may overestimate adiposity in children who have shorter statures or higher muscle mass and may underestimate adiposity in children with very low muscle mass. However, given its low cost, clinical utility and practicality, it is broadly used in clinical environments. It is therefore applied as an initial screen in assessing a patient's risk for obesity and obesity-related comorbidities. Due to the fact that children's BMI measurements change dramatically with age and differ with sex, age-and sex-specific BMI percentiles based on the Center for Disease Control (CDC) growth charts are used in place of raw BMI values[6]. Cutoff points for increased health risks are defined according to the 2007 expert committee recommendations convened by the department of Health and Human Services[5]. These guidelines suggest that a BMI of less than the 85^{th} percentile is unlikely to pose health risk, whereas a BMI greater than or equal to the 95^{th} percentile would confer significant risk. The terms "overweight" are therefore applied to a BMI $\geq 85^{th}$ percentile and "obesity" to a BMI $\geq$ the 95^{th} percentile. While the CDC

growth charts are useful for a large percentage of patients with overweight and obesity, BMI percentiles beyond the 97th percentile are not clinically useful, as large changes in BMI result in small percentile changes at the extreme. Therefore an additional metric, percentage of BMI at the 95th percentile (%BMIp95), is used to better assess and follow patients with severe obesity, defined as a BMI greater than or equal to 120% of the 95th percentile for age and sex[7].

2. What Does Pediatric Obesity and Overweight look like in the US?

On a population level, obesity disproportionately affects children from racial/ethnic minority backgrounds. African-American and Latino children display higher BMI scores from a young age and maintain a higher BMI growth trajectory compared to their non-Hispanic White counterparts [8]. According to some studies looking at disparities in obesity prevalence, obesity seems to emerge and is sustained earlier in Hispanic children relative to African Americans, but both groups experience higher BMIs by the 8th grade relative to non-Hispanic White children[9]. The morbidities associated obesity, such as hypertension and type II diabetes, are also disproportionately diagnosed in minority children and tend to be seen more in boys [10]. Having diseases such as elevated blood pressure or diabetes in childhood confers further risk of these diseases carrying on into adulthood and increases overall risk of mortality from cardiovascular or metabolic diseases [10, 11].

The risk factors associated with obesity are complex and intertwined. In general, poverty is positively associated with obesity prevalence[5]. There is evidence that genes play a role in obesity risk, and having one or both parents with obesity, increases the risk of a child developing obesity significantly [12]. However, the rapid increase in

prevalence at a population level suggests that environmental factors play a greater role than genetic shifts in the population[11]. Many associations with obesity risk such as infant birth weight, increased screen time, sleep patterns, and neighborhood-level factors have been described, but their interdependence and individual contribution to a patient's risk are largely undefined, making prediction, and prevention particularly difficult[6, 13-16].

Childhood obesity and overweight have shown to be predictors of future obesity, putting patients at risk for the eventual development of obesity-related comorbidities.[17] The medical complications of obesity are far reaching and include a range of life altering disorders including hypertension, diabetes mellitus, non-alcoholic fatty liver disease, dyslipidemia, asthma, and sleep apnea[18]. Managing these co-morbidities incur significant cost to individual patients and healthcare systems. One study estimates the lifetime cost for elementary students aged 6-11 with obesity to be $31,869 for boys and $39,815 for girls due mainly to the care required for comorbidity management [19].

3. Current Guidelines on Addressing Pediatric Obesity

In 2007, an expert committee was formed to revise the 1998 recommendations on childhood obesity. The recommendations were rooted in the latest evidence-based data and the experience of clinical experts to address prevention, assessment, and treatment of childhood overweight and obesity. The guidelines suggest that all children ages 2 years and older be screened with initial BMI measurements, family history of obesity and obesity-related disorders, and current diet and lifestyle practices. If a patient has a BMI that is $\geq 85^{th}$ percentile, the first steps a provider should take are to assess the medical and behavioral risks of the individual patient. Medical risk assessment includes screening for

common comorbid conditions such as hypertension, type 2 diabetes, hyperlipidemia, and non-alcoholic fatty liver disease. It was recommended by the committee that laboratory tests to screen for and diagnose such conditions be conducted every 2 years for children ages 10 years and older with obesity (or with overweight if they have associated risk factors) [5]. Behavioral assessment includes identifying obesogenic behaviors such as elevated screen time, fast food consumption, sugar-sweetened beverage intake, and sedentary lifestyle. Providers should then take steps to address overweight and obesity, and the guidelines make suggestions of four different treatment stages. These stages are: stage 1 prevention plus, stage 2 structured weight management, stage 3 comprehensive multidisciplinary approach and stage 4 tertiary care intervention. Each stage builds from office-based counseling for lifestyle and family recommendations (stage 1) to nutrition and psychological counseling (stages 2 and 3). Stage 4 uses interventions such as medications, very low calorie diets, and bariatric surgery[5]. In cases of a child not reaching a desired weight goal or in the presence of significant comorbidities, pharmacotherapy can be considered. Orlistat is the only FDA approved medication for the treatment of overweight and obesity in adolescents. Moderate improvements in BMI have been associated with the use of Orlistat however, unlike in adult counterparts, improvement in lipids or insulin sensitivity have not been consistently shown. Metformin has also shown some ability to improve BMI in some short-term obesity studies when used in conjunction with lifestyle modifications. Reported results on lipid and insulin sensitivity have been variable and Metformin is not FDA approved for weight reduction in pediatric patients [20].

4. Current Practice vs. Guidelines

Pediatric primary care providers (PCPs) are the cornerstone of addressing pediatric obesity as many successful interventions rely on PCPs to screen for and manage children with elevated BMI [21, 22]. Studies report that patients and families see their primary care provider as a reliable source of information and their recommendations have positive impacts on weight management [23, 24]. However, suboptimal rates of diagnosis of overweight and obesity based on BMI percentile in pediatric primary care persist [25, 26]. One 2011 study based on self-reported practice, found that less than 50% of primary care providers assessed BMI regularly in children and 58% reported rarely, or only sometimes using BMI percentiles to track weight [27]. Another 2011 study, found that pediatric providers reported unfamiliarity with the 2007 practice guidelines and diagnostic criteria for overweight and obesity suggesting that uptake of new practices has been slow[28]. Use of the Electronic Health Record (EHR) imparts the ability to auto-calculate BMI percentiles theoretically improving provider attention and diagnosis. Yet, despite some improvement with broad implementations of the EHR, children with overweight and obesity are still underdiagnosed[25, 29]. Counseling behaviors amongst providers have also been shown to be variable depending on factors such as sex, personal beliefs and attitudes[25]. In particular, younger children (2-5 years old) and children with overweight are more likely to be underdiagnosed, not receive diet and exercise counseling and have an absence of screening studies [2, 25, 30]. Perceived barriers to providing adequate care are often reported to be the sensitivity of the topic, clinic time constraints, and feelings of futility [3, 31, 32]. These inconsistencies across providers present missed opportunities to engage with families early, influence BMI trajectories, and provide high-quality care.

5. Methods of Assessing Provider Attention to Pediatric Weight Status

Given the growing need for PCP attention to childhood obesity and the suboptimal rates of diagnosis and screening, it is important to identify methods to support clinicians in this task. Broad use of the EHR puts researchers in a position to easily collect large amounts of data regarding physician practice. While manual chart review is still widely done, the process is laborious and may often limit sample sizes. Electronic phenotyping involves automated identification of subjects based on exclusion and inclusion criteria present in stored clinical data. Electronic phenotyping is typically used to identify patients with certain characteristics for a given purpose i.e.; a clinical trial, or retrospective study. In a study published in 2018 titled, "Algorithm to detect pediatric provider attention to high BMI and associated medical risk," Dr. Christy Turer and colleagues developed an electronic phenotyping algorithm using extractable EHR variables to indicate adherence to the 2007 expert committee guidelines on childhood obesity. Using diagnostic codes, laboratory studies, referrals, medications and procedures they categorized provider behavior in response to elevated BMI measurements into one of three phenotypes: "no attention", attention to "BMI Alone", and attention to "BMI/Medical Risk". Validation of the performance of the electronic phenotypes using manual chart review showed excellent sensitivity and specificity to detect provider attention types in pediatric clinics in Dallas, Texas[33]. By employing an algorithm to evaluate clinician behavior, Turer et al created a tool that went beyond identifying patients with disease characteristics to identifying encounters that follow guideline-based care. Furthermore, a follow up study published in 2019 by Turer et. al, demonstrated that children categorized by the algorithm as having primary care visits with attention to

elevated BMI and/or obesity-related medical risk were more likely to have improvement in weight status at follow-up visits [34]. Based on the results of these studies out of Texas and the anticipated benefits of using electronic phenotypes to augment provider practices, we sought to replicate and externally validate the Turer algorithm[33] in the Yale New Haven pediatric primary care setting.

STATEMENT OF PURPOSE

Hypothesis: We anticipate that the algorithm developed by Turer et. al 2018 to identify attention to elevated BMI and weight-related comorbidities among pediatric primary care clinicians would be applicable to 6-12 year-old children with overweight/obesity, defined as a BMI $\geq 85^{th}$ percentile, seen for well child visits in Yale New Haven Health pediatric primary care practices. Specifically, we hypothesize that implementation of this algorithm among patients at Yale-affiliated practices will yield a sensitivity and specificity for predicting attention (per manual review of EHR documentation) that are similar to the Turer study at a health system in Dallas, Texas. We also anticipate, based on data from previous research, that children with obesity or severe obesity will be more likely to be assigned an attention category in comparison to children with overweight, and non-Hispanic Black and Hispanic children will be more likely to be assigned an attention category than their non-Hispanic White counterparts [2, 25, 30, 35]. We expect to see variations on clinical practice based on trainee level as has been documented previously and therefore predict that children with encounters in the summer months (July- September), when new physicians begin their residency training, will be more likely to receive "no attention" than children seen later in other months [28, 36, 37]. Lastly, we predict that children with public versus private health care payors will be more likely to receive higher levels of attention.

Specific Aim 1: To externally validate the algorithm described by Turer et. al 2018 among 6-12 year old children with overweight/obesity seen at the Yale New Haven Hospital-affiliated pediatric primary care practices.

Specific Aim 2: To examine associations between pediatric provider attention and 1) weight category (overweight, obesity, severe obesity), 2) insurance type, 3) race/ethnicity and 4) season of encounter.

METHODS

Data Source:

We examined the records of 300 randomly selected patients ages 6-12 with two or more measurements of elevated BMI percentiles (≥85th percentile) for age and sex who were seen for well child visits on two or more occasions at any one of three pediatric primary care practices in the Yale New Haven Health system: the Yale Pediatric Primary Care Clinic (PCC), Yale Health Center (YHC), and Saint Raphael Campus Primary Care Clinic (SRC) from June 1, 2018 to May 31, 2019. If multiple encounters for the same patient occurred within that time period, we examined the encounter from the first chronological date. We categorized children's weight status based on the Center for Disease Control (CDC) growth charts from 2000 which classifies BMI-for-age into the following categories stratified by sex: overweight ≥ 85th to < 95th percentile, and obesity ≥95th percentile[38]. The intention of this study was to only examine patients with overweight or obesity.

Exclusion Criteria:

Patients were excluded from the study if they had less than two recorded BMI measurements above the 85th percentile to ensure that we were not examining visits with aberrant or incorrect BMI recordings. We also excluded children that were taking medications or had conditions that impact growth and nutrition (e.g., pregnancy, thyroid dysfunction, growth hormone abnormalities and sex hormone abnormalities).

Measures and Data Collection:

We extracted the following variables from the medical records of eligible patients:

a. Visit and problem list diagnosis codes entered on the date of the encounter

b. Referrals entered on the date of the encounter

c. Procedures/ lab orders entered on the date of the encounter

d. Medication lists queried for prescriptions written on day of the encounter

e. Age calculated in months based off of patient's birthdate and age at visit

f. BMI calculated using height and weight on the date of the visit

g. BMI categorization defined as overweight ($\geq 85^{th}$ – < 95^{th} percentile), obese ($\geq 95^{th}$ - <120% of the 95^{th} percentile) and severely obese (>120% of the 95^{th} percentile) using CDC growth charts BMI for age.

h. Sex (Male or Female)

i. Race/ethnicity defined as non-Hispanic Black, non- Hispanic White, Hispanic and Asian.

j. Insurance type defined as public (Medicaid), private (Blue Cross Blue Shield, Managed, or other commercial insurance), and Uninsured (self-pay or missing)

k. Provider type: defined as Nurse Practitioner, Physician Assistant, Physician (Attendings and Fellows), and Resident

Construction and Implementation of Algorithm to Detect Provider Attention:

The algorithm was modeled after the electronic phenotype described by Turer et. al[33]. After collecting the diagnostic codes, laboratory studies, medications, procedures and referrals used by our cohort, patient visits were classified into the following broad attention types: No Attention, Attention to BMI alone, Attention to Comorbidities alone,

and Attention to BMI and Comorbidities. The cohort was then sub-classified into comorbidity subtypes (Attention to Diabetes, Attention to Fatty Liver Disease, Attention to Hyperlipidemia and Attention to Vitamin D Deficiency) based on criteria listed by Turer et al (Figure 1). Criteria for classifying visits into attention types are listed in Table 1 for broad categories and Table 2 for comorbidity sub-types. The criteria were defined by reviewing the diagnostic codes, laboratory studies, medications, referrals and procedures used by Turer's team and comparing them to the corresponding values used in our population. Given that the original study was conducted using ICD-9 diagnostic codes, we first converted all codes into ICD-10 using the following website: http://www.icd10codesearch.com/.

Attention Type	Diagnosis Code (ICD-10)	Medicines and Laboratory Studies	Referrals and Procedures
BMI alone	E66.09 Obesity due to excess calories, unspecified obesity severity E66.09, Z68.54 Obesity due to excess calories without serious comorbidity with body mass index (BMI) in 95th to 98th percentile for age in pediatric patient E66.3 Overweight E66.3, Z68.53 Overweight peds (BMI 85-94.9 percentile) E66.3, Z68.54 Body mass index (BMI) of 95th to 99th percentile for age in pediatric patient E66.9 Obesity (BMI 30-39.9)/Obesity, unspecified classification, unspecified obesity type, unspecified whether serious comorbidity present E66.9, Z68.53 Obesity, pediatric, BMI 85th to less than 95th percentile for age E66.9, Z68.54 Obesity peds (BMI >=95 percentile)/BMI (body mass index), pediatric 95-99% for age, obese child structured weight management/multidisciplinary intervention category R63.3 Feeding difficulties R63.5 Weight gain/abnormal weight gain Z68.41BMI 40.0-44.9, adult (HC Code) Z68.53 BMI (body mass index), pediatric, 85% to less than 95% for age Z68.54 BMI pediatric, greater than or equal to 95% for age Z71.3 Nutritional counseling Z72.4 Inappropriate diet and eating habits	Orlistat DNA methylation analysis for Angelman or Prader Willi Syndrome Mutation analysis for Angelman or Prader Willi Syndrome *No patients in our population were taking medicines to treat obesity *No patients in our population had lab studies for Angelman, Prader Willi syndrome targeted gene mutation analysis or DNA methylation analysis DNA methylation analysis	Nutrition counseling Exercise counseling Nutrition Referral
Attention to BMI and Comorbidity	≥1 diagnosis code for BMI and ≥1 diagnosis code for: • Elevated blood pressure/hypertension • Acanthosis, prediabetes, diabetes type 2 • Lipid disorders • Fatty liver disease • Vitamin D deficiency *See table 2 for codes broken up by comorbidity	Labs to screen for diabetes, lipid disorders, fatty liver disease, or vitamin D deficiency (2 lab orders required, because increased likelihood lab ordered to screen for comorbidities related to overweight/obesity) • Medicines to treat hypertension, diabetes, lipid disorders, or vitamin D deficiency *See table 2 for laboratory studies broken up by comorbidity	Referral to tertiary weight management clinic, other medical specialist (endocrinologist, GI, Cardiologist, etc) *See table 2 for referrals broken up by comorbidity
Attention to Comorbidity Alone	≥1 diagnosis code for: • Elevated blood pressure/hypertension • Acanthosis, prediabetes, diabetes type 2 • Lipid disorders • Fatty liver disease • Vitamin D deficiency *See table 2 for codes broken up by comorbidity	Labs to screen for diabetes, lipid disorders, fatty liver disease, or vitamin D deficiency (2 lab orders required, because increased likelihood lab ordered to screen for comorbidities related to overweight/obesity) • Medicines to treat hypertension, diabetes, lipid disorders, or vitamin D deficiency *See table 2 for laboratory studies broken up by comorbidity	Referral to tertiary weight management clinic, other medical specialist (endocrinologist, GI, Cardiologist, etc) *See table 2 for referrals broken up by comorbidity
No attention	None of the above	None of the above	None of the above

Table 1. Diagnostic codes, laboratory studies, referrals and procedures that were used to qualify for each attention type based on the data from Yale pediatric primary care practices

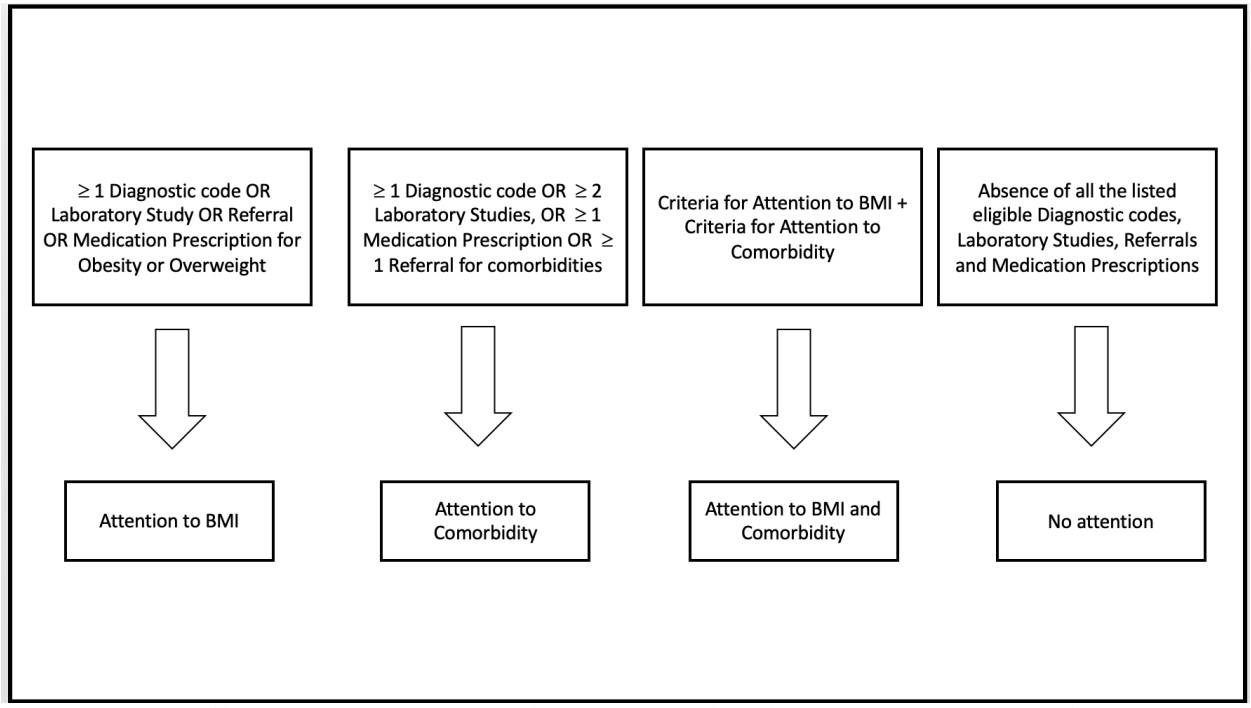

Figure 1: Overview of how attention types are assigned based on EHR collected variables. Please see Table 1 for specific diagnostic codes, laboratory orders, medication prescriptions and referrals for each attention type.

Comorbidity Attention Type	Diagnosis Codes	Referrals	Laboratory Studies	Medicines
Attention to Hypertension	R03.0: Elevated BP without diagnosis of hypertension/single episode of elevated blood pressure reading/ borderline hypertension	Pediatric Nephrology	*No lab orders were used in original study for this comorbidity	acetazolamide amlodipine atenolol hydralazine propranolol
Attention to Diabetes	E11.65: Uncontrolled type 2 diabetes mellitus without complication, without long-term current use of insulin L83: Acanthosis Nigricans R73.03: Attention to Diabetes R73.09: Elevated hemoglobin A1c Z13.1: Diabetes mellitus screening	Endocrinology, Diabetes & Metabolism	Glucose Glucose, fasting Glucose, gray top HEMOGLOBINA1C Insulin, total (BH GH LMW Q YH) Insulin, total (BH GH Q YH)	insulin aspar prot-insulin aspart insulin aspart insulin degludec insulin detemir insulin glargine insulin lispro insulin lispro protamine-lispro metformin metformin ER
Attention to Hyperlipidemia	Z13.220: Screening for cholesterol level	Pediatric cardiology Cardiovascular Disease	Cholesterol, Total HDL cholesterol LDL CHOLESTEROL, DIRECT LIPID PANEL LIPID PANEL WITH LDL/HDL RATIO (L)	atorvastatin

			Lipid profile w/ non-HDL cholesterol (L)	
Attention to Fatty Liver Disease	R74.0: Elevated ALT measurement R74.8: Elevated liver enzymes R94.5: Elevated LFTs	Pediatric Hepatology	ALT ALT+AST AST Hepatic function panel Hepatic function panel (LFT) LIVER FUNCTION TESTS (YH) Comprehensive metabolic panel Comprehensive metabolic panel without glucose	
Attention to Vitamin D Deficiency	E55.9: Vitamin D deficiency		QuestAssureD 25-OH vitamin D, (D2,D3), LC/MS/MS (LMW Q) QuestAssureD 25-OH vitamin D, (D2,D3), LC/MS/MS (BH LMW Q) VITAMIN D, 25-HYDROXY, TOTAL (GH L) Vitamin D 25 hydroxy (BH LMW) Vitamin D, 25-hydroxy, LC/MS/MS (Q YH)	cholecalciferol (vitamin D3) ergocalciferol

Table 2: Values used to qualify for attention to each comorbidity subtype based on data from our Yale associated practices.

Reference Standard:

To evaluate the performance of the electronic phenotype, an independent chart review process was done to validate the algorithm's ability to detect "No Attention". The dual purpose of the review was to manually examine the values used by the algorithm (diagnostic codes, labs, etc.) and to inquiry the visit encounter for other evidence of attention for which the algorithm was not designed to detect such as written text, or media entries. This process was done with substantial input from Dr. Christy Turer and attempts were made to maintain fidelity between the chart review she conducted and ours.

We began by examining the chart review/abstraction guide and questionnaire used in the Turer et al 2018 study to review 300 charts. We first converted her questionnaire from a paper copy to an online version in Qualtrics™ survey software (Qualtrics, Provo, UT). Several questions in the original questionnaire were intended to collect data for

other projects and were thus eliminated from our survey. Given that there are institutional differences in EPIC layout and note templates, our chart review guide had to be adapted to give clear directions for how to locate the desired information. In conjunction with Dr. Turer and in effort to replicate her team's process as much as possible, we also modified and added questions to the chart abstraction questionnaire to improve clarity and completeness. Examples of this include looking for laboratory orders in addition to laboratory results and adding a question to manually look at visit and problem list diagnostic codes associated with an encounter. Chart reviews included reviewing the growth chart, problem list, medication list, laboratory studies, family history, externally uploaded media, and visit notes associated with the visit date for each patient. Each chart was reviewed systematically using the Qualtrics survey. We used two separate reviewers (AG, AF) to examine 30 charts (10% of total) and compared responses to each Qualtrics survey question. Discrepancies in the responses were resolved with either a third-party reviewer (MS) or direct discussion between the reviewers. Interrater reliability was measured using the kappa statistic and interpreted using the guidelines outlined by Koch and Landis[39].

A difference in our chart review process in comparison to the original project was the selection of charts. Dr. Turer's team randomly selected 100 charts from each attention type (No Attention, Attention to BMI, Attention to Comorbidity). To control for bias, we chose to blindly review 300 charts from the cohort of 6-12 year olds and compare assigned attention types after completion of the review.

Statistical Analysis:

Our primary outcome was the algorithm's sensitivity and specificity of predicting no attention versus any attention compared to the reference standard.

The secondary analysis looked at demographic differences between attention types (no attention, attention to BMI alone, attention to Comorbidities alone, and attention to BMI and Comorbidities), for both chart review and algorithm, using Chi-squared tests of association.

Kappa statistics were computed by hand for interrater reliability (n = 2) of the chart review. All other analyses was completed using SAS version 9.4 (SAS Institute, Cary, NC).

Responsibilities:

The thesis primary author (AGB) was responsible for IRB writing and approval (with oversight from Dr. Sharifi), data randomization, creation of the chart review tool and review of clinic encounters. Kaitlin Maciejewski, MS (biostatistician) developed the SAS code to implement the algorithm and to conduct the statistical analysis. Additional support clarifying which variables were included in the original algorithm in Texas and general guidance was provided by Christy Boling Turer MD, MHS. Ada Fenick, MD assisted with the duplicate review of 10% of clinical encounters and helped refine the chart review tool together with Drs. Turer and Sharifi.

RESULTS

Demographics:

We reviewed 329 charts to identify 300 encounters that met our inclusion criteria and excluded 29 charts due to BMI measurements not meeting inclusion criteria. The mean±SD age of the sample was 10±1.87 years and 58.3% of children were male. Table 3 displays the demographics and encounter characteristics of the sample. In terms of weight categorization, 15.3% met criteria for severe obesity defined as ≥120% of 95th percentile, 41.3% met criteria for obesity, and 43.3% met criteria for overweight. The most prevalent race/ethnicity was Hispanic/Latino, compromising 42.7% of the cohort, Non-Hispanic Black was the next most prevalent 31% followed by Non-Hispanic White (15%) and Asian/Other (11.3%). Of the clinics included in our cohort, the majority (63.7 %) of encounters were conducted at the PCC, 24.7% at YHC, and 11.7% at SRC. The majority of patients in our sample had a public insurance payor type (64.7%) and the remainder had a private payor (14.3%) such as a managed healthcare or "Blue-cross Blue-shield," or other means (20.7%). Of the visit encounters examined, 49% were conducted by resident physicians with attending supervision, 25% by nurse Practitioners, 17.3% by attendings and 8.3% by physician assistants.

Table 5 displays the prevalence of attention to BMI and/or obesity-related comorbidities stratified by demographic and encounter characteristics. We observed statistically significant differences in the likelihood of classification as "no attention" by BMI category ($p<0.001$ for chart review and $p<0.001$ for algorithm) and by clinician type ($p<0.001$ for chart review and $p<0.001$ for algorithm). We did not observe statistically significant differences by race/ethnicity, season of encounter or insurance type.

Validation of the Algorithm:

Of the 30 charts comprehensively reviewed by two reviewers (AG and AF), kappa scores suggested substantial inter-observer agreement: 0.697 (95% confidence interval 0.482-0.912). Of the 300 charts reviewed in total, 50 were assigned no attention by chart review compared to 99 assigned no attention by the algorithm. Chart review identified 102 charts as attention to BMI, 4 as attention to comorbidity alone, and 144 as attention to BMI and Comorbidity. The algorithm correctly identified 66 and 82 as Attention to BMI and Attention to BMI and Comorbidity, respectively. This yielded a sensitivity of 94.0%, specificity of 79.2%, positive predictive value 47.5%, and negative predictive value 98.5% (Table 4).

A review of the charts for which the algorithm incorrectly labeled the encounter as "no attention," the discordance between the algorithm and chart review was due to evidence of attention in the form of free text within the progress note. Of the 52 encounters incorrectly labeled as no attention" by the algorithm, 83% had evidence of documentation in the assessment and plan and 49% had evidence of attention in the subjective sections of the progress note. A few encounters were incorrectly labeled as attention by the algorithm due to the use of qualifying laboratory studies, diagnosis codes, or referrals but not in the context of weight management. For example, in one encounter the provider ordered screening lipids as part of routine care during a 10 year-old's well child visit in addition to vitamin D screening as part of deficiency screening in refugee clinic; these two lab orders satisfy the algorithm's categorization of "attention to comorbidity" but were appropriately not identified as attention during chart review. These findings are summarized in Figure 2.

Characteristic	*Overall Sample N (%)*
BMI categorization	
Overweight	130 (43.3%)
Obesity (Class 1 only)	124 (41.3%)
Severe Obesity (Class 2 and 3)	46 (15.3%)
Race/ethnicity	
Asian/Other	33 (11.0%)
Hispanic or Latino	128 (42.7%)
Non-Hispanic Black	94 (31.3%)
Non-Hispanic White	45 (15.0%)
Provider type	
Attending	53 (17.7%)
Nurse Practitioner	75 (25.0%)
Physician Assistant	25 (8.3%)
Resident	147 (49.0%)
Season	
Fall	72 (24.0%)
Spring	62 (20.7%)
Summer	106 (35.3%)
Winter	60 (20.0%)
Insurance type	
Private Health care	44 (14.7%)
Public health care	194 (64.7%)
Other / missing	62 (20.7%)
Clinic	
SRC	35 (11.7%)
YHC (pediatrics)	74 (24.7%)
YNH-PCC	191 (63.7%)

Table 3 Demographics of the sample of 300 encounters examined. Obesity class 1 was defined as BMI percentile ≥ 95th to <120%BMIP95. Obesity Class 2 and 3 was defined as BMI percentile ≥120 %BMIp95.

Chart review attention type	Algorithm flag				
	Attention to BMI	**Attention to BMI and Comorbidity**	**Comorbidities only**	**No attention**	**Total**
Attention to BMI	66	2	0	34	102
Attention to BMI and Comorbidity	27	82	18	17	144
Comorbidities only	0	0	3	1	4
No attention	0	0	3	47	50
Total	93	84	24	99	300

Table 4: Comparison of algorithm and chart review attention assignments

Characteristic	Chart-Review Defined			Algorithm Defined		
	Attention to BMI or BMI & comorbidity or Comorbidity alone	**No Attention**	**P value***	**Attention to BMI or BMI & comorbidity or Comorbidity alone**	**No Attention**	**P value**
BMI categorization						
Overweight	95 (38.0%)	35 (70.0%)	<0.001	70 (34.8%)	60 (60.6%)	<0.001
Obesity (class 1 only)	110 (44.0%)	14 (28.0%)		93 (46.3%)	31 (31.3%)	
Severe Obesity (class 2 and 3)	45 (18.0%)	1 (2.0%)		38 (18.9%)	8 (8.1%)	
Race/ethnicity						
Asian/Other	30 (12.0%)	3 (6.0%)	0.19	25 (12.4%)	8 (8.1%)	0.47
Hispanic or Latino	108 (43.2%)	20 (40.0%)		88 (43.8%)	40 (40.4%)	
Non-Hispanic Black	79 (31.6%)	15 (30.0%)		58 (28.9%)	36 (36.4%)	
Non-Hispanic White	33 (13.2%)	12 (24.0%)		30 (14.9%)	15 (15.2%)	
Provider type						
Attending	50 (20.0%)	3 (6.0%)	<0.001	40 (19.9%)	13 (13.1%)	<0.001
Nurse Practitioner	52 (20.8%)	23 (46.0%)		40 (19.9%)	35 (35.4%)	
Physician Assistant	8 (3.2%)	17 (34.0%)		5 (2.5%)	20 (20.2%)	
Resident	140 (56.0%)	7 (14.0%)		116 (57.7%)	31 (31.3%)	
Season						
Fall	61 (24.4%)	11 (22.0%)	0.75	46 (22.9%)	26 (26.3%)	0.83
Spring	54 (21.6%)	8 (16.0%)		40 (19.9%)	22 (22.2%)	
Summer	86 (34.4%)	20 (40.0%)		73 (36.3%)	33 (33.3%)	
Winter	49 (19.6%)	11 (22.0%)		42 (20.9%)	18 (18.2%)	
Insurance type						
Private Health care	38 (15.2%)	6 (12.0%)	0.68	30 (14.9%)	14 (14.1%)	0.06
Public health care	159 (63.6%)	35 (70.0%)		122 (60.7%)	72 (72.7%)	
Other / missing	53 (21.2%)	9 (18.0%)		49 (24.4%)	13 (13.1%)	
Clinic						
SRC	16 (6.4%)	19 (38.0%)	<0.001	12 (6.0%)	23 (23.2%)	<0.001
YHC (pediatrics)	63 (25.2%)	11 (22.0%)		53 (26.4%)	21 (21.2%)	
YNH-PCC	171 (68.4%)	20 (40.0%)		136 (67.7%)	55 (55.6%)	

Table 5: Chi-squared test of association between receiving provider attention and relevant clinical and demographic variables

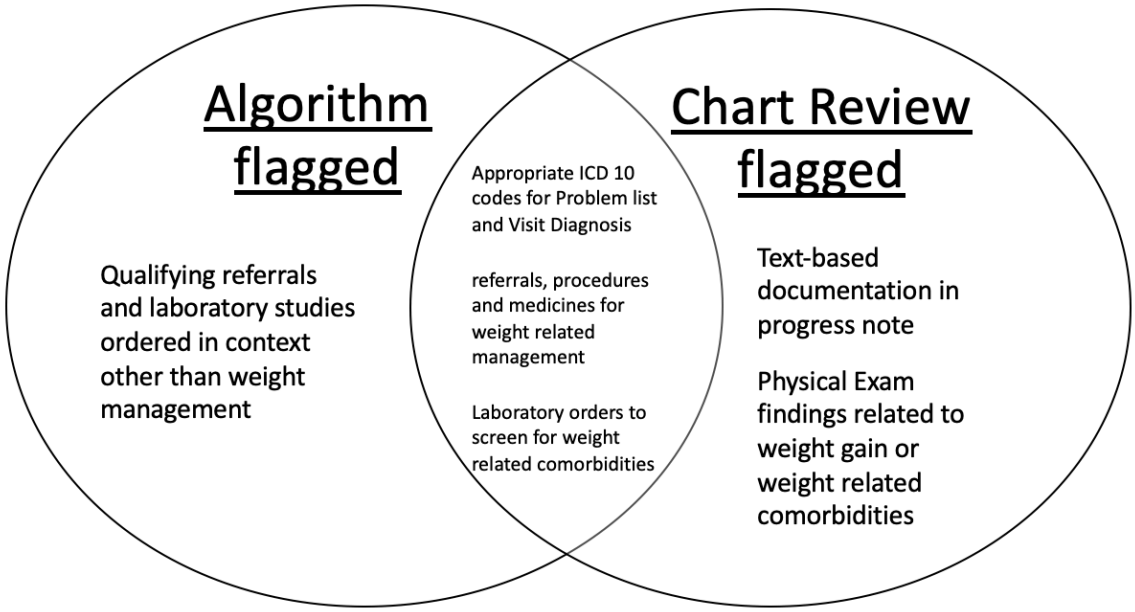

Figure 2: Summary of discrepancies between algorithm flags and chart review flags.

DISCUSSION

The purpose of this project was to assess the external validity of an electronic phenotyping algorithm that evaluates clinician attention to elevated BMI and associated comorbidities during pediatric primary care visits against a reference standard of manual chart review. Our chart review revealed that 250 out of 300 encounters (83%) had some evidence of attention to weight status or weight-related comorbidities in the electronic record associated with the visit, and 50 of the 300 (16%) encounters examined lacked evidence of any attention. The electronic phenotyping algorithm flagged 201 out of 300 (67%) visits with an attention type (attention to BMI, attention to comorbidity or attention to BMI and comorbidity) and 99 out of 300 (33%) visits with no evidence of attention. The application of the Turer algorithm to the Yale pediatric primary care setting had a sensitivity of 94.0% and a specificity of 79.2% for predicting "no attention" versus any attention type. The positive predictive value of "no attention" relative to any attention type in our cohort was 47.5% and the negative predictive value of "no attention" was 98.5%. When examining the encounters labeled as "no attention", we found significant differences in assignment based on weight category, consistent with previous research [2, 25, 30], as well as by clinician type.

In comparison with the performance of the phenotyping algorithm in the original Turer et al study population, our algorithm had a slightly lower sensitivity and substantially lower specificity for identifying "no attention". Using the sensitivity and specificity from the Turer et. al, study we would expect a positive predictive value of 79.0% in our sample. However, our minimally modified algorithm yielded a positive predictive value of 47.5%. The suboptimal specificity and positive predictive value could

be explained by a lack of generalizability of the original algorithm. The algorithm developed by Turer et al was refined based on the practices of documentation conducted at the UT Southwestern clinic sites. Several iterations of code were reviewed to enhance the performance of the algorithm in their practices. However, documentation practices regarding the structured inputs that the algorithm evaluated such as problem list entry or ICD 10 codes vary between sites. An example of this could be seen in the ordering of Vitamin D labs and the reporting of Vitamin D Deficiency. Vitamin D deficiency has been known to be associated with overweight/obesity and was included in the Turer algorithm as an indicator of attention to weight status. However, Vitamin D deficiency is routinely evaluated in the relatively large population of refugee patients seen at the Yale PCC, regardless of weight status, causing the algorithm to incorrectly identify attention to a comorbidity (false positives) and limiting the generalizability of the algorithm to the Yale primary care setting.

Although our electronic phenotype had a lower sensitivity and specificity than the original study, its use of multiple EHR inputs make it a better predictor of attention than ICD diagnosis codes alone. For example, compared to a study completed at Yale pediatric primary care sites between November 2011 and May 2015, only 11% of children with overweight had a visit diagnoses of overweight in their medical record, 37% of children with obesity has a diagnosis code, and 54% of children with severe obesity received had the corresponding diagnosis code [35]. Other studies examining the use of diagnosis codes for overweight and obesity, have shown similarly low rates[2, 40, 41]. This suggests that while diagnosis of obesity or overweight may still be sub-optimal, clinicians may be using other areas of the EHR to denote attention to weight status and

weight related comorbidities in their patients. Evaluating provider behavior solely on diagnosis codes provides an incomplete assessment of the clinic encounter.

Although a substantial improvement over diagnosis codes alone, the Turer algorithm is limited to structured data and does not utilize the large amount of unstructured data available in EHR encounter, i.e, free text in clinical notes. It has already been established that clinicians have suboptimal rates of entering diagnosis codes and problem list entries. However, the clinical progress note is an important medical-legal document that many providers rely on to communicate information. We found that of the encounters that were incorrectly labeled as no attention" by the algorithm, a large percentage of them (83%) had evidence of documentation in the assessment and plan and/or in the subjective sections (49%) of the progress note. By excluding free-text elements in the clinical progress note from our evaluation of a clinical encounter, we are capturing an incomplete picture of the encounter. Future studies looking to continue use of electronic phenotyping to assess provider behavior should include natural language processing to capture unstructured EHR inputs.

Another limitation of this study is that it primarily evaluated the algorithm in mostly academic primary care settings. Although one of our study sites was a non-teaching clinical environment, it is still difficult to predict performance in different settings such as private practices.

In conclusion, our findings suggest that implementation of an electronic phenotyping algorithm without adaption for the local site may result in lower specificity than originally reported out of the health system in which the algorithm was developed. Although still clinically useful and superior to other available options, adaptions such as

the use of natural language processing could enhance the precision and accuracy of this phenotyping algorithm as a pragmatic tool to detect attention to elevated BMI and associated comorbidities.

REFERENCES

1. Center for Disease Control. *Childhood obesity facts*. 2018, January 29; Available from: https://www.cdc.gov/healthyschools/obesity/facts.htm.
2. Sharifi, M., et al., *Evaluating the implementation of expert committee recommendations for obesity assessment.* Clin Pediatr (Phila), 2013. **52**(2): p. 131-8.
3. Shreve, M., A. Scott, and K. Vowell Johnson, *Adequately Addressing Pediatric Obesity: Challenges Faced by Primary Care Providers.* South Med J, 2017. **110**(7): p. 486-490.
4. Savinon, C., et al., *Childhood obesity: Can electronic medical records customized with clinical practice guidelines improve screening and diagnosis?* J Am Acad Nurse Pract, 2012. **24**(8): p. 463-71.
5. Barlow, S.E. and C. Expert, *Expert committee recommendations regarding the prevention, assessment, and treatment of child and adolescent overweight and obesity: summary report.* Pediatrics, 2007. **120 Suppl 4**: p. S164-92.
6. Yanovski, J.A., *Pediatric obesity. An introduction.* Appetite, 2015. **93**: p. 3-12.
7. Gulati, A.K., D.W. Kaplan, and S.R. Daniels, *Clinical tracking of severely obese children: a new growth chart.* Pediatrics, 2012. **130**(6): p. 1136-40.
8. Guerrero, A.D., et al., *Racial and Ethnic Disparities in Early Childhood Obesity: Growth Trajectories in Body Mass Index.* J Racial Ethn Health Disparities, 2016. **3**(1): p. 129-37.
9. Rendall, M.S., et al., *Hispanic and black US children's paths to high adolescent obesity prevalence.* Pediatr Obes, 2012. **7**(6): p. 423-35.
10. Flynn, J.T., et al., *Clinical Practice Guideline for Screening and Management of High Blood Pressure in Children and Adolescents.* Pediatrics, 2017. **140**(3).
11. Berenson, G.S. and g. Bogalusa Heart Study, *Health consequences of obesity.* Pediatr Blood Cancer, 2012. **58**(1): p. 117-21.
12. Kumar, S. and A.S. Kelly, *Review of Childhood Obesity: From Epidemiology, Etiology, and Comorbidities to Clinical Assessment and Treatment.* Mayo Clin Proc, 2017. **92**(2): p. 251-265.
13. Wethington, H., L. Pan, and B. Sherry, *The association of screen time, television in the bedroom, and obesity among school-aged youth: 2007 National Survey of Children's Health.* J Sch Health, 2013. **83**(8): p. 573-81.
14. Miller, A.L., J.C. Lumeng, and M.K. LeBourgeois, *Sleep patterns and obesity in childhood.* Curr Opin Endocrinol Diabetes Obes, 2015. **22**(1): p. 41-7.
15. Yang, Y., et al., *A cross-sectional study of the influence of neighborhood environment on childhood overweight and obesity: Variation by age, gender, and environment characteristics.* Prev Med, 2018. **108**: p. 23-28.
16. Weihrauch-Bluher, S. and S. Wiegand, *Risk Factors and Implications of Childhood Obesity.* Curr Obes Rep, 2018. **7**(4): p. 254-259.
17. Ward, Z.J., et al., *Simulation of Growth Trajectories of Childhood Obesity into Adulthood.* N Engl J Med, 2017. **377**(22): p. 2145-2153.
18. Pulgaron, E.R., *Childhood obesity: a review of increased risk for physical and psychological comorbidities.* Clin Ther, 2013. **35**(1): p. A18-32.

19. Sonntag, D., *Why Early Prevention of Childhood Obesity Is More Than a Medical Concern: A Health Economic Approach.* Ann Nutr Metab, 2017. **70**(3): p. 175-178.
20. Matson, K.L. and R.M. Fallon, *Treatment of obesity in children and adolescents.* J Pediatr Pharmacol Ther, 2012. **17**(1): p. 45-57.
21. Taveras, E.M., et al., *Comparative Effectiveness of Clinical-Community Childhood Obesity Interventions: A Randomized Clinical Trial.* JAMA Pediatr, 2017. **171**(8): p. e171325.
22. Sharifi, M., et al., *Accelerating progress in reducing childhood obesity disparities: exploring best practices of positive outliers.* J Health Care Poor Underserved, 2013. **24**(2 Suppl): p. 193-9.
23. Ackard, D.M. and D. Neumark-Sztainer, *Health care information sources for adolescents: age and gender differences on use, concerns, and needs.* J Adolesc Health, 2001. **29**(3): p. 170-6.
24. Bean, M.K., et al., *Impact of motivational interviewing on outcomes of an adolescent obesity treatment: results from the MI Values randomized controlled pilot trial.* Clin Obes, 2018. **8**(5): p. 323-326.
25. Reyes, I., *An Evaluation of the Identification and Management of Overweight and Obesity in a Pediatric Clinic.* J Pediatr Health Care, 2015. **29**(5): p. e9-14.
26. Klein, J.D., et al., *Adoption of body mass index guidelines for screening and counseling in pediatric practice.* Pediatrics, 2010. **125**(2): p. 265-72.
27. Huang, T.T., et al., *Pediatricians' and family physicians' weight-related care of children in the U.S.* Am J Prev Med, 2011. **41**(1): p. 24-32.
28. Rausch, J.C., E.R. Perito, and P. Hametz, *Obesity prevention, screening, and treatment: practices of pediatric providers since the 2007 expert committee recommendations.* Clin Pediatr (Phila), 2011. **50**(5): p. 434-41.
29. Imoisili, O.E., et al., *Screening and Referral for Childhood Obesity: Adherence to the U.S. Preventive Services Task Force Recommendation.* Am J Prev Med, 2019. **56**(2): p. 179-186.
30. Camp, N.L., et al., *Identification of Overweight and Obesity in Low-Income Minority Children by Pediatric Providers and Child Characteristics Associated With Underrecognition.* J Pediatr Health Care, 2019. **33**(2): p. 162-168.
31. Yarborough, B.J., et al., *Responding to pediatric providers' perceived barriers to adolescent weight management.* Clin Pediatr (Phila), 2012. **51**(11): p. 1063-70.
32. Busch, A.M., A. Hubka, and B.A. Lynch, *Primary Care Provider Knowledge and Practice Patterns Regarding Childhood Obesity.* J Pediatr Health Care, 2018. **32**(6): p. 557-563.
33. Turer, C.B., C.S. Skinner, and S.E. Barlow, *Algorithm to detect pediatric provider attention to high BMI and associated medical risk.* J Am Med Inform Assoc, 2018.
34. Turer, C.B., et al., *Association of Clinician Behaviors and Weight Change in School-Aged Children.* Am J Prev Med, 2019. **57**(3): p. 384-393.
35. Lydecker, J.A. and C.M. Grilo, *The Missed Diagnosis and Misdiagnosis of Pediatric Obesity.* Psychother Psychosom, 2017. **86**(3): p. 173-174.
36. Cyr, P.R., et al., *Weighty Problems: Predictors of Family Physicians Documenting Overweight and Obesity.* Fam Med, 2016. **48**(3): p. 217-21.

37. Morais, A., et al., *Characteristics of Correctly Identified Pediatric Obesity and Overweight Status and Management in an Academic General Pediatric Clinic.* Clin Pediatr (Phila), 2018. **57**(10): p. 1168-1175.
38. Kuczmarski, R.J., et al., *CDC growth charts: United States.* Adv Data, 2000(314): p. 1-27.
39. Landis, J.R. and G.G. Koch, *The measurement of observer agreement for categorical data.* Biometrics, 1977. **33**(1): p. 159-74.
40. Young, E.L., *Increasing Diagnosis and Treatment of Overweight and Obese Pediatric Patients.* Clin Pediatr (Phila), 2015. **54**(14): p. 1359-65.
41. Reed, M., et al., *Identification, Prevention, and Management of Childhood Overweight and Obesity in a Pediatric Primary Care Center.* Clin Pediatr (Phila), 2016. **55**(9): p. 860-6.