Penn Healthy Diet survey: pilot validation and scoring

Charlene W. Compher[1]*, Ryan Quinn[2], Richard Haslam[3], Elizabeth Bader[4], Joellen Weaver[5], Scott Dudek[6], Marylyn D. Ritchie[6], James D. Lewis[7]† and Gary D. Wu[7]†

[1]*University of Pennsylvania School of Nursing, Department of Biobehavioral Health Science, Philadelphia, PA, USA*
[2]*University of Pennsylvania School of Nursing, Department of Biostatistics, Philadelphia, PA, USA*
[3]*University of Dublin, School of Medicine, Dublin, Republic of Ireland*
[4]*University of Pennsylvania, Philadelphia, PA, USA*
[5]*University of Pennsylvania Health System, Penn Medicine Biobank, Philadelphia, PA, USA*
[6]*Department of Genetics and Institute for Biomedical Informatics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA*
[7]*University of Pennsylvania Perelman School of Medicine, Division of Gastroenterology and Hepatology, Philadelphia, PA, USA*

Abstract

Though diet quality is widely recognised as linked to risk of chronic disease, health systems have been challenged to find a user-friendly, efficient way to obtain information about diet. The Penn Healthy Diet (PHD) survey was designed to fill this void. The purposes of this pilot project were to assess the patient experience with the PHD, to validate the accuracy of the PHD against related items in a diet recall and to explore scoring algorithms with relationship to the Healthy Eating Index (HEI)-2015 computed from the recall data. A convenience sample of participants in the Penn Health BioBank was surveyed with the PHD, the Automated Self-Administered 24-hour recall (ASA24) and experience questions. Kappa scores and Spearman correlations were used to compare related questions in the PHD to the ASA24. Numerical scoring, regression tree and weighted regressions were computed for scoring. Participants assessed the PHD as easy to use and were willing to repeat the survey at least annually. The three scoring algorithms were strongly associated with HEI-2015 scores using National Health and Nutrition Examination Survey 2017–2018 data from which the PHD was developed and moderately associated with the pilot replication data. The PHD is acceptable to participants and at least moderately correlated with the HEI-2015. Further validation in a larger sample will enable the selection of the strongest scoring approach.

Keywords: Diet: HEI-2015: ASA24: Biobank: Survey: Screener: Score: NHANES

Dietary patterns of food intake have been linked to the most common noncommunicable diseases in developed countries[1–5]. Most clinical care utilises electronic medical record systems that represent a trove of information that can be leveraged for quality improvement and research. Unfortunately, many electronic medical records contain very limited nutrition intake information[6], data that would be particularly helpful not only for nutrition counseling but also for nutrition epidemiology research in clinical contexts.

Self-reported recalls of the previous day's diet are commonly used to assess food intake and diet quality. The validated Automated Self-Administered 24-hour (ASA24®) is an online tool which allows individuals to self-report their dietary intake using either a computer or a smartphone[7]. The automated multi-pass method used requires multiple queries about the list of the previous day's foods, amounts eaten, food preparation methods, condiments or seasoning added and details forgotten, a cognitively demanding process. It is estimated that most individuals are able to complete the ASA24 in less than 30 min[8]. While self-administered 24-h dietary recalls are an advance for research[7,9], challenges remain. Individuals who never learned to type, who do not work with computers routinely or who have limited internet access may need to complete the ASA24 using their smartphone. If they are visually challenged, hurried or have limited cell phone minutes or web access, the quality of their dietary recall information may be

compromised or incomplete. Moreover, the amount of time required for completion and data analysis may lead to selection bias in research and generally make use of self-administered 24-h dietary recalls impractical for clinical care. To assess food patterns on a larger scale with lower participant burden, a simpler solution is needed.

Dietary screening tools hold promise for filling this gap in nutrition information when care is provided in time-limited settings[10–16]. Screeners can obtain information about key aspects of the diet quickly, typically in less than 5 min[17,18]. The Penn Healthy Diet (PHD) screening tool[19] was developed using dietary intake data in the 2017–2018 National Health and Nutrition Examination Survey (NHANES)[20] with a focus on food groups that comprise the Healthy Eating Index (HEI)-2015[21,22]. To better reflect the diversifying USA population, since 2011 NHANES has oversampled several groups including Hispanics, Non-Hispanic Blacks, Non-Hispanic Asians, Non-Hispanic Whites, persons over age 80 years and those living below 185 % of the federal poverty line[20]. The PHD queries the number of times (0 to 5 or more) items from twenty-nine food groups were consumed during the previous day. The request to report number of exposures was selected in lieu of portion sizes that are both difficult to assess and cognitively challenging to estimate. Commonly reported foods in the NHANES 2017–2018 diet recall data were added to each food group to provide currently available examples in the USA food supply. Items delineating types of protein foods and behavioural items were added to aid in facilitation of nutrition counseling goals.

A simple scoring algorithm identified twelve food groups (fruit juice, fruit, green/leafy vegetables, red/orange vegetables, whole grains, milk, seafood, plant proteins, nuts/seeds, sugary beverages, refined grains and cheese) and three behavioural items (full fat dairy, added butter/gravy and added oil) that were moderately or strongly associated with HEI-2015 variables. The simple scoring approach with PHD data correlated strongly (Spearman's rho = 0·75) with the HEI-2015 score from NHANES diet recall information[19]. However, in the initial study, participant responses to the PHD screener questions were simulated based on their reported food intake in the NHANES 24-h dietary recalls. Further comparison of the PHD against related items in a concurrently collected 24-h recall is needed to confirm these initial findings and to refine the scoring approach.

The purposes of this research were to assess participant experience with the PHD screener in a clinical sample, to validate the accuracy of the PHD screening tool against related items in the ASA24 and to explore new scoring algorithms with relationship to the Healthy Eating Index-2015 computed from the ASA24 data. We hypothesised that the PHD screener would be acceptable to participants, that the screener would identify the consumption of similar foods to those identified with the ASA24 and that a scoring algorithm based on data collection via the PHD would correlate with parameters in the HEI generated from ASA24 data.

Materials and methods

Participants

The Penn Medicine BioBank (PMBB) has enrolled 174 712 racially diverse participants (56 % female; 17 % Black, 71 %

White, 4 % Asian, 3 % Other and 6 % Unknown) with a median of seven years of prospective health and disease data mapped to ICD diagnostic codes as well as the complete electronic medical record data (imaging, clinical laboratory measures, procedures, clinical notes, etc.) for each participant[23]. The PMBB also includes a biorepository of blood and tissue samples for genetic and other omics assays; to date, whole exome sequence and genome-wide genotype data are available for approximately 45 000 participants. However, dietary intake data are not collected in any systematic way. The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of University of Pennsylvania (protocol codes: 808346 approved 07/01/2008, 813913 approved 4/3/2013 and 817977 approved 6/6/2013) for studies involving humans. Potential respondents provided informed consent by their electronic response to the approved email invitation.

Design

A convenience sample of 385 PMBB participants who were recently treated by the gastroenterology service was invited to participate by email or telephone call. Email invitations were sent via the secure Research Electronic Data Capture (REDCap) application[24], with a follow-up reminder email one week later, if a patient did not respond to the initial invitation. The email invitation led to the PHD screener survey and a unique login credential and password for the ASA24 website to complete the dietary recall. Participants who did not have an email on record with the PMBB were contacted by telephone to offer the opportunity to participate by email invitation, mail to their home address or verbally by telephone. If a patient did not respond to the telephone call, a voicemail message was left which detailed the participation and remuneration process. Patient responses were recorded in the REDCap survey system. If a patient requested telephone assistance to complete the ASA24, questions regarding the patient's previous day's dietary intake were asked in the order and format as they appear in the ASA24 system. The patient's answers were entered directly into the ASA24 site as they responded to the questions. Participants were incentivised $5 for completing the PHD and $20 for completing the ASA24, using the Greenphire® ClinCard program. The incentive amounts were designed to recognise the likely time commitment of the subject to complete the surveys.

Participant experience questions were included at the end of the screener to gauge the ease of participation, and the device that was used to complete the survey and willingness to complete the PHD again. Time to complete the screener was computed by subtracting start time from completion time in REDCap. The ASA24 site reports time to complete the ASA24.

Participant demographics were abstracted from the electronic medical record. Diagnosis codes (ICD-10) were collected and used to calculate Charlson Comorbidity Index[25,26].

Statistical analyses

Summary statistics were computed for participants' clinical and demographic characteristics using means and standard deviations and frequencies and percentages for continuous and

categorical measures respectively. Likert-scale responses to the patient experience questions to the PHD were described as frequency and percentage.

The PHD screener queries the number of times (0 to 5+) on the previous day that foods from twenty-nine food groups were eaten (water, coffee/tea, fruit juice, whole fruit, whole grains, refined grains, alcohol, sugary beverages, diet soda, milk, yogurt, cheese, eggs, poultry, seafood, plant proteins, red meat/pork, cured meat, fried foods, fast food meals, nuts, desserts, sweet snacks and salty snacks)[19]. Similarly to the HEI-2015, the PHD asks for the number of servings of unsweetened fruit juice. However, fruit drink and fruit punch are included in the sugary beverages question to capture this significant source of added sugar and processed foods. The PHD also includes seven behavioural questions with significant information for diet counseling (adding energy content to coffee/tea such as sweeteners or cream; adding artificial sweeteners; using full fat dairy products; adding salt at the table; adding butter/gravy and using oils in cooking). While the PHD was designed based on the HEI-2015, the HEI-2015 does not address intake of alcohol, water or coffee/tea, or the behavioural variables in the PHD in a discrete way. These items were included in the PHD to aid in diet counseling for adequate fluid intake or potential sources of energy content or salt.

The simple scoring approach originally developed for the PHD was based on those food groups that had at least a moderate association with HEI-2015 variables[19]. The score added one point for each time the previous day healthy foods were reported (fruit juice; fruit; green/leafy vegetables; red/orange vegetables; whole grains; milk; fish/seafood; plant proteins and nuts/seeds) and reverse scored those with negative correlations (sugary beverages; refined grains and cheese). Two behavioural questions added one point for a negative response (added butter/gravy; full fat dairy) and one for a positive response (added oil). The score range was 0–63 with higher score indicating a healthier diet.

The food group file provided by the ASA24 comprises self-reported serving size of foods ingested by the respondent. By contrast, the PHD asks only how many times a food was eaten the previous day, without consideration of serving size. Each food item reported by participants in the ASA-24 diet recall was coded by a nutrition science expert into categories which correspond to each of the twenty-nine screener food group items. The three questions about eating behaviours in the simple scoring approach were not included because no equivalent item was found in the ASA24 food list. While these behaviours have recognised impact on health and intake of energy and saturated fat, they are not recognisable as discrete variables in the HEI-2015 score.

To evaluate the similarity of related responses to the PHD relative to the ASA24, two approaches were used based on the presence or absence of a food in the ASA24 report. For each food group, observations were coded as *insertions* if the food was reported in the screener but not in the diet recall. Observations were coded as *deletions* if the food was reported in the diet recall but not in the screener. Lastly, observations were coded as *congruent* when a food group was reported consistently in both the diet recall and the screener. Frequency tables were produced to summarise insertions, deletions and congruent reporting across each food group. Cohen's Kappa statistic was computed to assess the extent of agreement in the presence and absence of each food group's reporting across the screener and diet recall responses. Cohen's Kappa assesses the extent to which two categorical ratings agree beyond chance, where negative scores indicate less than chance agreement and scores 0–1 indicate greater than chance agreement[27]. Data published as kappa results were interpreted as follows: Values ≤ 0 as no agreement, 0–0·20 as no to slight agreement, 0·21–0·40 as fair, 0·41–0·60 as moderate, 0·61–0·80 as substantial and 0·81–1·00 as almost perfect agreement[28,29].

A screener score was computed using the simple scoring algorithm previously published[19] with exclusion of the three behavioural variables, using Spearman correlation coefficients to assess associations between screener data and the HEI 2015 total score and subcomponents. Correlation coefficients were interpreted as follows: Values 0–0·2, negligible; 0·2–0·39 weak; 0·4–0·60 moderate; 0·7–0·89 strong and 0·9–1·0 very strong correlation[30].

Two additional novel data-driven approaches to development of a screener scoring algorithm were taken utilising simulated screener responses derived from NHANES 2017–2018 data as the training set but then testing the strength of the models in the 3P study data. Using the NHANES data as a training dataset, simulated screener responses from the NHANES participants were operationalised as predictors of the HEI-2015 score in a regression tree model. The regression tree model was selected based on Spearman correlations between simulated screener values and HEI-2015 scores. The model utilised reduced error pruning and employed a minimum leaf size of ten and maximum tree depth of 9 to avoid potential overfitting to the training dataset. The selected regression tree model was then operationalised as a scoring algorithm to compute predicted HEI-2015 scores for the 3P study data. Utilising the 3P study data, the Spearman correlation between the PHD screener score and the ASA-24-derived HEI-2015 score was then computed to determine concurrent validation of the scoring algorithm. Regression tree modelling was conducted using Proc HPSPLIT in SAS 9·4.

In this model, each simulated screener response (excluding dichotomous behavioural questions) was included as a predictor of the HEI-2015 scores. Regression coefficients derived from the multivariable model were then applied as item weights to compute a screener score for both the NHANES and 3P study data. Spearman correlations were computed to assess scoring algorithm performance on the NHANES data, as well as to determine concurrent validation using the 3P study data. Statistical analyses were conducted using SAS 9·4.

Results

Participants

A total of 385 individuals were invited to participate in the pilot survey. Participants were contacted by email (*n* 355) if an email address was available or by telephone call (*n* 33). For two participants contacted by telephone who were interested in completing the PHD screener by telephone, each question and possible response was read out in entirety. One patient

requested to complete the ASA24 recall with assistance over the telephone. The overall response rate was 60 (15·6 %) and 23 (5·9 %) of the total invited continued to complete the ASA24. However, another 27 (45 %) never logged in, and 5 (8 %) logged in but quit before completing the dietary recall.

The demographic characteristics of respondents to the surveys are shown in Table 1. The mean age was 59 years, 75 % were female, predominantly non-Hispanic White and obese. The Charlson Comorbidity Index score was 1·32.

Patient experience with the Penn Healthy Diet

Responses to the patient experience questions and response time are shown in Table 2. All respondents reported that the PHD was easy to complete, and 99 % would be willing to repeat the survey. The most common frequency of willingness to complete was monthly, but all would complete it at least annually. The most common device used to complete the PHD was the cell phone (52 %) followed by the computer (39 %) and tablet (5 %). The median times to completion of the screener were 5 min for cell and tablet and 7·5 min for the computer. While the ASA24 site does not identify which device was used to complete the dietary recall, the minutes required for completion was median (IQR) 21 (13–39) min.

Accuracy of Penn healthy diet screener item responses relative to dietary recall

Comparison of the accuracy of the PHD relative to responses to the same patient's response to the ASA24 is shown in Table 3. Congruent responses occurred far more commonly than insertions or deletions. Only three items had < 70 % congruence. The whole grain bread item appeared on the PHD before white bread, and 52 % of responses reported whole grain bread on the PHD but not the ASA24. By contrast, 44 % of respondents reported white bread on the ASA24 but not on the PHD. Nuts were reported congruently by 65 %, as insertions on the PHD by 30 % and as deletions by 4 %. In assessing statistical agreement of food categories in the screener and diet recall, Kappa statistics ranged from −0·12 to 0·88 (mean = 0·44) with 15/29 screener items exhibiting moderate to strong agreement with the ASA24.

Screener scoring approaches

Using the simple scoring algorithm recently published[19], in respondents to both the PHD and the ASA24, the PHD score was 14·5±4·68 and the HEI-2015 score based on ASA24 data was 52·7 ±16·5. The screener score exhibited a moderate correlation with the HEI-2015 total score ($r = 0.59$, $P = 0.0034$). Online Supplementary Fig. 1 presents a heatmap which visualises associations between screener items and total score $v.$ HEI-2015 subcomponents and total score. As with the HEI-2015, a higher screener score implies a diet with greater nutrient than energy density.

The regression tree model selected the same twelve variables as the simple scoring algorithm, but added nine additional variables (alcohol, desserts, cured meat, red meat/pork, fast food meals, coffee/tea, poultry, savory snacks and eggs).

The multivariable regression model constructed using NHANES 2017–2018 data with simulated screener responses also selected the

Table 1. Demographic characteristics of survey respondents

	Respondents to PHD (n 60)		Respondents to PHD and ASA24 (n 23)	
Variable	n	%	n	%
Age, years				
Mean	59·05		54·39	
SD	13·8		16·87	
Sex				
Female	45	75	17	73·9
Male	15	25	6	26·1
Race				
Black	11	18·3	5	21·7
White	48	80	18	78·3
Other	2	3·3	1	4·3
Ethnicity				
Hispanic	2	3·3	0	
Non-Hispanic	57	95	23	100
Not reported	1	1·7	1	4·3
BMI (kg/m²)				
Mean	32·50		35·04	
SD	6·33		1·13	
18·5–24·9	7	11·6	1	4·3
25–29·9	17	28·3	2	8·7
30 or more	36	60	20	87
Charlson Comorbidity Index				
Mean	1·32		1	
SD	1·58		1·13	

Table 2. Patient experience with Penn healthy diet screener

Question	Response	%	Median	Interquartile range
How difficult was the screener?	Somewhat easy	21		
	Very easy	79		
Would you repeat the screener?	Yes	98·4		
	No	1·6		
If you would repeat the screener, how often would you repeat it?	Annually	5		
	Biannually	11·5		
	Monthly	50·8		
	Weekly	32·8		
What device did you use to complete the screener?	Cell phone	52·4		
	Computer	39·3		
	Tablet	4·9		
	Other	3·3		
Measured time to complete screener, minutes	Cell		5	4–7
	Tablet		5	4–823
	Computer		7·5	5–12
	Other		n/a	

same twelve variables as the simple approach but added the same nine additional variables as the regression tree but also diet soda and yogurt (online Supplementary Table 4). The five most important food groups with negative betas were refined grains ($\beta = -2.19 + 0.19$), cured meat ($\beta = -2.19 + 0.23$), fast food meals ($\beta = -1.14 + 0.32$), cheese ($\beta = -1.11 + 0.14$) and sweetened beverages ($\beta = -0.84 + 0.19$), while those with positive beta coefficients were whole fruit($\beta = 3.79 + 0.23$), yogurt ($\beta = 3.54 + 0.70$), dark green vegetables ($\beta = 2.98 + 0.82$), whole grains ($\beta = 2.84 + 0.18$) and nuts/seeds ($\beta = 2.33 + 0.16$).

Table 4 contains the correlation coefficients comparing responses to the PHD score to the HEI-2015 score from dietary

Table 3. Comparison of Penn healthy diet screener responses to ASA-24 responses

Variable	Reported on Penn Healthy Diet Screener n (%)		Reported on Automated Self Administered 24 Recall		Reported in both Penn Healthy Diet Screener and Diet Recall		Insertions Reported in Penn Healthy Diet Screener but not Diet Recall		Deletions Reported in Diet Recall but not in Penn Healthy Diet Screener		Kappa statistic
	n	%	n	%	n	%	n	%	n	%	
Alcohol	8	34·8	6	26·1	21	91·3	2	8·70	–		0·80
Desserts	5	21·7	8	34·8	16	69·6	2	8·70	5	21·7	0·26
Cheese	17	73·9	14	60·9	16	69·6	5	21·7	2	8·70	0·32
Salty Snacks	12	52·2	9	39·1	16	69·6	5	21·7	2	8·70	0·40
Coffee	16	69·6	16	69·6	21	91·3	1	4·35	1	4·35	0·79
Cold Cuts	7	30·4	7	30·4	21	91·3	1	4·35	1	4·35	0·79
Cookies	10	43·5	9	39·1	18	78·3	3	13·0	2	8·70	0·55
Cured Meat	6	26·1	4	17·4	21	91·3	2	8·70	–		0·75
Diet Soda	5	21·7	1	4·35	19	82·6	4	17·4	–		0·28
Egg	6	26·1	5	21·7	22	95·7	1	4·35	–		0·88
Fast Food Meals	1	4·35	2	8·70	20	87·0	1	4·35	2	8·70	−0·06
Fish	3	13·0	3	13·0	19	82·6	2	8·70	2	8·70	0·23
Fried Foods	2	8·70	5	21·7	18	78·3	1	4·35	4	17·4	0·18
Green Vegetables	17	73·9	15	65·2	17	73·9	4	17·4	2	8·70	0·39
Fruit Juice	5	21·7	1	4·35	19	82·6	4	17·4	–		0·28
Beef, Pork	10	45·5	9	39·1	19	82·6	2	8·70	1	4·35	0·72
Milk	7	30·4	3	13·0	19	82·6	4	17·4	–		0·51
Nuts	13	56·5	7	30·4	15	65·2	7	30·4	1	4·35	0·34
Red/Orange Vegetables	11	47·8	16	69·6	16	69·6	1	4·35	6	26·1	0·40
Poultry	8	34·8	8	34·8	19	82·6	2	8·70	2	8·70	0·62
Sugary Beverages	6	26·1	4	17·4	17	73·9	4	17·4	2	8·70	0·24
Plant Proteins	6	26·1	2	8·70	19	82·6	4	17·4	–		0·43
Water	23	100	18	78·3	18	78·3	5	21·7	–		.
White Bread	6	27·3	16	69·6	11	47·8	1	4·35	10	43·5	0·14
White Rice	12	52·2	12	52·2	19	82·6	2	8·70	2	8·70	0·65
Whole Fruit	14	60·9	14	60·9	19	82·6	2	8·70	2	8·70	0·63
Whole Grain Bread	14	60·9	2	8·70	11	47·8	12	52·2	–		0·12
Whole Grain Cereals	3	13·0	2	8·70	18	78·3	3	13·0	2	8·70	−0·12
Yogurt	10	43·5	4	17·4	17	73·9	6	26·1	–		0·43

Table 4. Spearman correlation coefficients between Penn Healthy Diet survey scores and Healthy Eating Index-2015 scores among training and validation data sets

PHD Scoring Method	Training Data: NHANES 2017–2018*	Validation Data: 3P Study Data**
Simple Algorithm	$r = 0.75$, $P < 0.0001$	$r = 0.59$, $P = 0.0034$
Regression Tree	$r = 0.80$, $P < 0.0001$	$r = 0.49$, $P = 0.0190$
Regression Weight	$r = 0.81$, $P < 0.0001$	$r = 0.67$, $P = 0.0004$

* Training data consists of simulated Penn Healthy Diet scores derived from 24 h diet Recall, and Healthy Eating Index-2015 score derived from 24 h diet recall.
** Validation data consist of Penn Healthy Diet scores derived from participant PHD responses, and Healthy Eating Index-2015 score derived from Automated Self Administered Dietary Assessment Tool 24 diet recall.

recalls using each of the three scoring approaches applied to the NHANES 2017–2018 training set and the 3P pilot validation set. Application of these three scoring approaches to the 3P data yielded lower regression coefficients than the NHANES 2017–2018 data (simple approach, 0·59 *v.* 0·75; regression tree, 0·49 *v.* 0·80 and regression weight $r = 0.67$ *v.* 0·81, respectively). Because the screener scores are of different magnitudes, a comparison of the three approaches statistically is not possible.

Discussion

There is substantial need for more efficient and user-friendly methods to gauge diet quality. We developed the PHD screener to help meet this need[19]. In this study, we demonstrated that the patient experience of the PHD screener was generally positive, and the time to completion was in the range expected for a dietary screener. Responses of times foods were eaten were predominantly congruent with the ASA24 data. All three PHD scoring approaches identified similar food groups, and the correlations based on 3P replication data were only moderate while those based on the NHANES derivation set were strong.

This PHD screener bears some similarities to other tools developed for use in European clinical settings. The Mediterranean Diet Adherence screener was developed for use in a large Mediterranean diet intervention in Spain[15], recently validated for use in the UK[14] and was recently updated with an energy-restricted version[16]. While the PHD screening tool targets similar foods that are more commonly available in the USA diet, neither the PHD nor Mediterranean Diet Adherence Survey permits assessment of energy intake. However, the close alignment of the PHD with the HEI-2015, a tool that is energy-adjusted, suggests that higher scores with the PHD would indicate a diet that is more nutrient- than energy dense. A brief diet quality assessment tool for use in French-

speaking Quebec was developed relative to the Alternative Healthy Eating Index score[12], but it has not yet been piloted in healthcare settings. Two tools have been developed in Finland, one for use in adult primary care[11] and the other for dietary counseling by professionals without specific training in nutrition[13]. The PHD survey was designed to support dietary counseling and to give a sense of overall diet quality.

As with the other European and Canadian recent screener evaluations[10–16], a strength of the PHD is that it was tested here in a sample from the setting in which it would be useful. The favourable subject experience data are also encouraging about the feasibility of employing the tool more broadly. However, the relatively low response rate to our unsolicited email request to participate suggests that a different approach to obtaining data from biobank participants is indicated. The low response rate also carries the risk that responders were not fully representative of other invitees in terms of interest in diet or willingness to report detailed dietary information. We note that the participants who participated had similar self-reported race, but a larger proportion of females when compared with published data from the PMBB. Women have been described by others as more interested nutrition and healthy diet than men[31,32]. While these similarities are encouraging, we cannot evaluate how representative this sample is of general hospital samples in the USA. A strength of the 3P pilot project is that the same subjects responded to both the PHD and the ASA24 surveys. While we cannot determine whether the same device was used for the ASA24 as for the PHD, the completion time was considerably shorter with the PHD, as expected. The most common device used for the PHD and probably the ASA24 was the cellphone. If the requirement of multiple screen views for the ASA24 is tedious by cellphone, this may be why so many started the diet recall but stopped before completion. It is encouraging that the same food groups were identified with both novel approaches to scoring, those with positive beta were foods scored positively (adequacy items) and those with negative beta were foods scored negatively (to take in moderation) by the Healthy Eating Index-2015. A weakness in the original version of the PHD tested here was that respondents appear to have entered white bread as whole grain bread, possibly because the whole grain option appeared earlier in the survey. Since white grains are eaten more commonly than whole grains in the USA, we will list these options prior to the whole grain options in future surveys to avoid confusion in responses for subjects who are not clear about the meaning of whole grain bread. A limitation to this pilot project and to the strength of our conclusions about a scoring algorithm is the small sample size. A larger validation study is underway using the reordered white *v.* whole grain options that will improve our ability to choose the best-performing scoring approach while also assessing the value of the screener to predict metabolomic signatures. We will continue to use the original simple scoring algorithm until a stronger option is clear.

In conclusion, the PHD screener has demonstrated high levels of acceptance by participants. The PHD screener is strongly correlated with the HEI-2015 derived from 24-h dietary recalls. However, further research is needed to identify the optimal scoring algorithm for clinical and research purposes. Nonetheless, the routine use of electronic medical records creates the potential to automate delivery and scoring of the PHD within the context of routine care. As such, the PHD holds enormous promise for nutrition-focused clinical care and research.

Supplementary material

For supplementary material accompanying this paper visit https://doi.org/10.1017/S0007114523001642

References

1. English LK, Ard JD, Bailey RL, *et al.* (2021) Evaluation of dietary patterns and all-cause mortality: a systematic review. *JAMA Netw Open* **4**, e2122277.
2. Finicelli M, Squillaro T, Di Cristo F, *et al.* (2019) Metabolic syndrome, Mediterranean diet, and polyphenols: evidence and perspectives. *J Cell Physiol* **234**, 5807–5826.
3. Bovolini A, Garcia J, Andrade MA, *et al.* (2021) Metabolic syndrome pathophysiology and predisposing factors. *Int J Sports Med* **42**, 199–214.
4. Schwingshackl L, Bogensberger B & Hoffmann G (2018) Diet quality as assessed by the Healthy Eating Index, Alternate Healthy Eating Index, Dietary Approaches to Stop Hypertension Score, and Health Outcomes: an Updated

Systematic Review and Meta-Analysis of Cohort Studies. *J Acad Nutr Diet* **118**, 74–100.e111.

5. Cena H & Calder PC (2020) Defining a healthy diet: evidence for the role of contemporary dietary patterns in health and disease. *Nutrients* **12**, 334.

6. Vadiveloo M, Lichtenstein AH, Anderson C, *et al.* (2020) Rapid diet assessment screening tools for cardiovascular disease risk reduction across healthcare settings: a scientific statement from the American Heart Association. *Circ Cardiovasc Qual Outcomes* **13**, e000094.

7. Park Y, Dodd KW, Kipnis V, *et al.* (2018) Comparison of self-reported dietary intakes from the Automated Self-Administered 24-h recall, 4-d food records, and food-frequency questionnaires against recovery biomarkers. *Am J Clin Nutr* **107**, 80–93.

8. NCI (2022) Automated Self-Administered 24-Hour (ASA24®) Dietary Assessment Tool. https://epi.grants.cancer.gov/asa24/ (accessed December 2022).

9. Thompson FE, Dixit-Joshi S, Potischman N, *et al.* (2015) Comparison of interviewer-administered and automated self-administered 24-hour dietary recalls in 3 diverse integrated health systems. *Am J Epidemiol* **181**, 970–978.

10. Schroder H, Benitez Arciniega A, Soler C, *et al.* (2012) Validity of two short screeners for diet quality in time-limited settings. *Public Health Nutr* **15**, 618–626.

11. Hemio K, Polonen A, Ahonen K, *et al.* (2014) A simple tool for diet evaluation in primary health care: validation of a 16-item food intake questionnaire. *Int J Environ Res Public Health* **11**, 2683–2697.

12. Lafreniere J, Harrison S, Laurin D, *et al.* (2019) Development and validation of a Brief Diet Quality Assessment Tool in the French-speaking adults from Quebec. *Int J Behav Nutr Phys Act* **16**, 61.

13. Lindstrom J, Aittola K, Polonen A, *et al.* (2021) Formation and validation of the Healthy Diet Index (HDI) for evaluation of Diet Quality in healthcare. *Int J Environ Res Public Health* **18**, 2362.

14. Papadaki A, Johnson L, Toumpakari Z, *et al.* (2018) Validation of the English version of the 14-item Mediterranean Diet adherence screener of the PREDIMED study, in people at high cardiovascular risk in the UK. Nutrients 10, 138.

15. Schroder H, Fito M, Estruch R, *et al.* (2011) A short screener is valid for assessing Mediterranean diet adherence among older Spanish men and women. *J Nutr* **141**, 1140–1145.

16. Schroder H, Zomeno MD, Martinez-Gonzalez MA, *et al.* (2021) Validity of the energy-restricted Mediterranean Diet adherence screener. *Clin Nutr* **40**, 4971–4979.

17. Thompson FE, Midthune D, Kahle L, *et al.* (2017) Development and evaluation of the National Cancer Institute's Dietary Screener Questionnaire scoring algorithms. *J Nutr* **147**, 1226–1233.

18. Thompson FE & Byers T (1994) Dietary assessment resource manual. *J Nutr* **124**, 2245S–2317S.

19. Compher CW, Quinn R, Burke F, *et al.* (2022) Development of the Penn Healthy Diet screener with reference to adult dietary intake data from the National Health and Nutrition Examination Survey. *Nutr J* **21**, 70.

20. CDC (2017–2018) National Health and Nutrition Examination Survey. https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2017 (accessed October 2021).

21. Reedy J, Lerman JL, Krebs-Smith SM, *et al.* (2018) Evaluation of the Healthy Eating Index-2015. *J Acad Nutr Diet* **118**, 1622–1633.

22. Krebs-Smith SM, Pannucci TE, Subar AF, *et al.* (2018) Update of the Healthy Eating Index: HEI-2015. *J Acad Nutr Diet* **118**, 1591–1602.

23. Verma A, Damrauer SM, Naseer N, *et al.* (2022) The Penn Medicine BioBank: towards a genomics-enabled learning healthcare system to accelerate precision medicine in a diverse population. *J Pers Med* **12**, 1974.

24. Research Electronic Data Capture (REDCap). https://www.project-redcap.org/ (accessed August 2023).

25. D'Hoore W, Bouckaert A & Tilquin C (1996) Practical considerations on the use of the Charlson comorbidity index with administrative data bases. *J Clin Epidemiol* **49**, 1429–1433.

26. Charlson ME, Carrozzino D, Guidi J, *et al.* (2022) Charlson comorbidity index: a critical review of clinimetric properties. *Psychother Psychosom* **91**, 8–35.

27. Tang W, Hu J, Zhang H, *et al.* (2015) Kappa coefficient: a popular measure of rater agreement. *Shanghai Arch Psychiatry* **27**, 62–67.

28. McHugh ML (2012) Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* **22**, 276–282.

29. Altman DG & Ebscohost (1991) *Practical Statistics for Medical Research*, 1st ed. London; New York: Chapman and Hall.

30. Schober P, Boer C & Schwarte LA (2018) Correlation coefficients: appropriate use and interpretation. *Anesth Analg* **126**, 1763–1768.

31. Wardle J, Haase AM, Steptoe A, *et al.* (2004) Gender differences in food choice: the contribution of health beliefs and dieting. *Ann Behav Med* **27**, 107–116.

32. Barebring L, Palmqvist M, Winkvist A, *et al.* (2020) Gender differences in perceived food healthiness and food avoidance in a Swedish population-based survey: a cross sectional study. *Nutr J* **19**, 140.