The utility of household Grocery Purchase Quality Index scores as an individual diet quality metric

Haley Wynne Parker[1], Carolina de Araujo[1], Anne N. Thorndike[2,3] and Maya K. Vadiveloo[1]*

[1]*Department of Nutrition and Food Sciences, University of Rhode Island, Kingston, RI 02881, USA*
[2]*Department of Medicine, Division of General Internal Medicine, Massachusetts General Hospital, Boston, MA 02114, USA*
[3]*Department of Medicine, Harvard Medical School, Boston, MA 02115, USA*

Abstract

The Grocery Purchase Quality Index (GPQI) reflects concordance between household grocery purchases and US dietary recommendations. However, it is unclear whether GPQI scores calculated from partial purchasing records reflect individual-level diet quality. This secondary analysis of a 9-month randomised controlled trial examined concordance between the GPQI (range 0–75, scaled to 100) calculated from 3 months of loyalty-card linked partial ($\geq$50 %) household grocery purchasing data and individual-level Healthy Eating Index (HEI) scores at baseline and 3 months calculated from FFQ (n 209). Concordance was assessed with overall and demographic-stratified partially adjusted correlations; covariate-adjusted percentage score differences, cross-classification and weighted κ coefficients assessed concordance across GPQI tertiles (T). Participants were middle aged (55·4 (13·9) years), female (90·3 %), from non-smoking households (96·4 %) and without children (70·7 %). Mean GPQI (54·8 (9·1) %) scores were lower than HEI scores (baseline: 73·2 (9·1) %, 3 months: 72·4 (9·4) %) and moderately correlated (baseline r 0·41 v. 3 months r 0·31, $P < 0·001$). Correlations were stronger among participants with $\leq$ bachelor's degree, obesity and children. Scores showed moderate agreement ($\kappa = 0·25$); concordance was highest in T3. Participants with high (T3) v. low (T1) GPQI scores had 7·3–10·6 higher odds of having HEI scores >80 % at both time points. Household-level GPQI was moderately correlated with self-reported intake, indicating their promise for evaluating diet quality. Partial purchasing data appear to moderately reflect individual diet quality and may be useful in interventions monitoring changes in diet quality.

Key words: Grocery purchase quality: Healthy Eating Index: Diet quality: Grocery purchase data

Poor diet quality is a preventable chronic disease risk factor[1]. Multi-level interventions aiming to increase healthy food access and availability in various food environments (i.e. schools, communities, workplaces)[2–7] have the potential to improve diet quality, and such efforts may be enhanced through use of novel dietary assessment methods that complement existing measures of intake[8]. The home food environment, determined by micro- (e.g. taste preferences, food security) and macro-level factors (e.g. policy, marketing), substantially influences individual-level diet quality[9], making it a desirable target for interventions.

Measuring the home food environment is challenging as few tools are validated and a range of data sources (e.g. receipts, inventories, checklists, etc.) reduces comparability[10–12]. The Grocery Purchase Quality Index (GPQI) presents an opportunity to evaluate home food environments more accurately using objective food purchasing data to calculate ratios of actual v. recommended spending within food groups as a proxy for diet quality[13]. Similar to the commonly used Healthy Eating Index (HEI)[14], the

GPQI assesses adherence to the Dietary Guidelines for Americans, containing both adequacy and moderation components[13]. However, unlike the HEI, the GPQI relies on less detailed data to categorise foods into broader food groups without assessing nutrient components, making it a more flexible tool for evaluating household diet quality, and potentially individual-level intake. Despite relying on less detailed data, a recent study found a strong positive correlation (r 0·70) between GPQI and HEI scores both calculated with objective, household-level purchasing data[15], demonstrating that the GPQI validly assesses diet quality despite broader categorisation of foods.

To date, limited research has explored the relationship between purchase quality and intake quality. Notably, Appelhans *et al.*[16] found a moderate correlation between HEI scores calculated from complete grocery purchase records and 24-h recalls ($\rho c = 0·57$, $P < 0·0001$), suggesting that grocery purchasing data may be a useful proxy of intake quality. Further research is needed to evaluate associations between household

grocery purchase quality and individual-level diet quality when only a partial grocery purchasing record is available and determine how shopper characteristics influence the association.

Within the Smart Cart pilot randomised controlled trial, the present study assessed whether diet quality of the home food environment provided valid estimates of individual-level diet quality. The GPQI reflected purchase quality based on 3 months of partial (≥50 %) grocery purchasing data from one supermarket while the HEI assessed individual-level diet quality prior to the start of the study and at 3 months so that concordance between measures could be compared both prospectively and retrospectively. Concordance between the GPQI and HEI was evaluated with: (1) overall and subgroup partially adjusted correlations, (2) percentage score differences adjusting for confounding variables, (3) cross-classifying participants with high or low diet quality and (4) weighted κ coefficients assessing consistency in tertile assignments.

Methods

This secondary data analysis was completed within the Smart Cart study (registered at clinicaltrials.gov, NCT03748056: Targeted Food Incentives to Improve Diet Quality and Health Among Adults), a pilot randomised controlled cross-over trial examining the impact of individually targeted coupons on grocery purchase quality; study details have been published elsewhere[17]. The Smart Cart study was conducted according to the guidelines laid down in the Declaration of Helsinki, and all procedures involving human subjects were approved by the University of Rhode Island Institutional Review Board; #1240194. Written informed consent was obtained from all subjects.

Participants (*n* 224) who regularly grocery shopped at the participating medium-sized grocery store in Rhode Island were recruited between July and October 2018. The participating grocery store is a full-service, independent, locally owned specialty market featuring an in-house bakery and extensive selection of fresh prepared foods that are known for its healthy options and local produce. Eligible participants were ≥18 years, English-speaking, the primary grocery shopper in their household, completed ≥50 % of their grocery shopping at the participating store, had email for study communication, were not pregnant or planning to become pregnant during the study and were enrolled or willing to enroll in the store's loyalty programme.

Recruited participants completed a demographics and grocery shopping habits survey using RedCap and a validated online FFQ reflecting 3 months of individual-level intake; the online FFQ contained branching logic with complex skip patterns to reduce participant burden and therefore included a variable number of questions for each participant[18,19]. Participants were randomised to the intervention group which received targeted weekly coupons and nutrition education content (based on purchased and self-reported diet history) or the control group which received non-targeted nutrition education content and occasional generic coupons (i.e. $2 off any purchase). After the first 13-week intervention period, 3-month FFQ were administered during a 2–4 week washout period, after which experimental groups crossed over.

Loyalty cards were used to collect purchasing data, distribute coupons and provide a 5 % discount on grocery purchases. The analytic sample (*n* 209; online Supplementary Fig. S1) excluded fifteen participants from the full sample with no purchases in any study period (*n* 8) or total period spending <1 % (*n* 4) or >99 % (*n* 3) of the sample distribution. A sensitivity analysis was conducted eliminating participants who completed the baseline or 3-month FFQ in <10 min (*n* 11); results were unchanged, so these participants were included in the analytic sample.

Dietary variables

Household-level diet quality was measured using the GPQI-16 which has been validated against the HEI-15, displaying reasonable convergent and criterion validity[15]. GPQI scores were calculated with objective point-of-sale purchasing data from the participating store over a 3-month period (September–December 2018). The GPQI is scored out of 75 points and comprised of eleven components, including eight adequacy (e.g. fruit, vegetables, where intake is encouraged) and three moderation components (e.g. refined grains, sweets, where intake is discouraged). GPQI scores are derived by comparing percentage of total spending within each component to standardised expenditure shares, which reflect the United States Department of Agriculture Food Plans (i.e. guidance for cost-effective grocery purchasing that aligns with national nutrition guidelines). Brewster *et al.*[15] derived standardised expenditure shares by census-weighting age- and sex-specific Food Plan component expenditures and averaging across four income strata. To calculate GPQI scores, all purchased items were assigned one of thirteen mutually exclusive categories (eleven GPQI components, non-categorised food items (i.e. oils, coffee, prepared and mixed dishes), or non-food items). After excluding non-food items, total spending was summed for overall and component-specific food purchases. Component scores are based on the ratio of actual percentage spending to standardised expenditure shares. A greater percentage spending on adequacy components (and lesser percentage spending on moderation components) reflects greater adherence to the Dietary Guidelines for Americans and equates to higher GPQI scores.

Individual-level diet quality was measured using the extensively validated HEI-2010[20]. HEI-2010 scores were calculated from the participants' responses to a 3-month FFQ administered at enrollment and 3 months. The HEI is scored out of 100 points and comprised of twelve components, including nine where intake is encouraged (i.e. adequacy) and three where intake is discouraged (i.e. moderation)[21]. The HEI includes all eight of the GPQI adequacy components (i.e. total fruit, whole fruit, total vegetables, greens and beans, whole grains, dairy products, total protein foods and seafood and nuts/plant proteins) plus fatty acids. Both indexes have three adequacy components, including refined grains which is common to both indexes; the GPQI additionally contains processed meats and sweets and sodas, while the HEI contains Na and 'empty calories' (includes solid fats, added sugar and alcohol)[15,21].

Consistent with best practices for validity analyses[22], the 3-month FFQ recall period is temporally aligned with GPQI scores; complete data were available for 182 of the analytic sample at 3 months. Because data were available at both time points, this analysis was able to assess whether baseline HEI (HEI$_{BL}$; July–October 2018) was predictive of future GPQI scores and whether the GPQI validly reflected 3-month HEI scores (HEI$_{MID}$; December 2018–January 2019). For both measures (HEI and GPQI), higher total and component scores indicate a healthier diet. Due to different score ranges (GPQI = 0–75; HEI = 0–100), scores were scaled to 100 points for comparability. Estimated energy and macronutrient intakes were calculated from FFQ responses but since energy and macronutrient content of grocery purchases could not be estimated, analyses strictly focused on diet quality overall and component scores.

Statistical methods

Analyses were completed in SAS 9.4 (SAS Institute Inc.) and the α threshold for significance was set at $P < 0.05$, with Bonferroni and Tukey corrections used to adjust the α threshold for multiple comparisons[23]. GPQI scores were sorted in ascending order and participants were divided into tertiles (T), which roughly categorised participants into groups of low, medium and high grocery purchase quality. Descriptive statistics (including means and standard errors and numbers and percentages) were calculated for the overall sample and compared across GPQI tertiles with unadjusted generalised linear and logistic regression models. The analytic sample's representativeness of the full study sample was examined by comparing descriptive characteristics among those included $v.$ excluded at 3 months using t tests for continuous variables and χ^2 for categorical variables.

Several validity measures assessed concordance between GPQI and HEI scores including correlation coefficients, percentage score differences, cross-classification and weighted κ coefficients. Bonferroni-corrected Pearson correlations (partially adjusted for age and sex) characterised the strength and direction of association between continuous GPQI and HEI scores at baseline and 3 months, stratified by shopper- and household-level characteristics.

Covariate-adjusted generalised linear models compared percentage score differences (i.e. GPQI – HEI) and mean HEI$_{BL}$, HEI$_{MID}$ and GPQI scores across GPQI tertiles for total and nine component scores common to both indexes. Covariates were selected partly based on previous literature; therefore, sex and age were included in all adjusted models[24]. Based on R^2 and mean score differences when singly added to the model, additional individual- and household-level measures self-reported at baseline were considered for adjustment, including baseline experimental group assignment, educational attainment ($\leq$ or >bachelor's degree), annual pre-tax household income (< or $\geq$\$100 000), continuous BMI (self-reported weight and height), household size (i.e. the number of people groceries were purchased for; 1, 2, or >2 people), children in household (yes/no), smokers in household (yes/no) and whether or not (yes/no) participants also grocery shopped at farmer's markets, superstores or other grocery stores. The fully adjusted model was controlled for age, sex, BMI and income; Tukey correction was used for *post hoc* comparisons.

Cross-classification was used to determine whether participants with high or low GPQI scores (i.e. T3 or T1) would be similarly classified as having high or low HEI score (i.e. defined *a priori* as $\geq$80 or <60 points from previous literature[14]). Logistic regression models adjusted for age, sex, BMI and income compared the odds of obtaining high ($\geq$80 $v.$ <80) and low (<60 $v.$ $\geq$60) HEI scores within each GPQI tertile.

Lastly, weighted κ coefficients characterised agreement between GPQI and HEI tertile rankings. For interpretability and to count partial agreements, Fleiss–Cohen quadratic weights were used[25]. A weight of one reflecting perfect agreement was given to identical rankings (e.g. GPQI T3 and HEI T3), a weight of 0·75 reflected partial agreement (e.g. HEI T2 and GPQI T1) and a weight of zero reflected opposite assignments (e.g. GPQI T1 and HEI T3).

The Smart Cart Study was initially powered to detect a 3 % difference in GPQI scores across experimental groups in a cross-over study design. Therefore, 224 participants were recruited, resulting in a final sample of 200 predicting 10 % loss-to-follow-up[17]. For this secondary analysis, to obtain 80 % power and detect a medium effect size (r 0·3)[26], a sample size of eighty-four was needed. The analytic sample (n 209) was adequately powered for this analysis.

Results

Table 1 presents descriptive characteristics within the overall sample and by GPQI tertiles, where GPQI T1 has the least healthy grocery purchases (range 22·8–50·9) and GPQI T3 (range 58·5–79) has the healthiest purchases. Mean age (55·4 (13·9) years) decreased from 58·1 to 53·6 years across increasing GPQI tertiles ($P_{\text{for trend}} = 0·05$). Similarly, mean BMI (25·4 (4·6) kg/m²) decreased from 26·8 in T1 to 24·1 in T3 ($P_{\text{trend}} < 0·001$). Participants were predominately female (90·3 %), higher socio-economic status (49·7 % income <\$100 000; 51·0 % $\leq$bachelor's degree), shopping for households with 2·3 (1·1) members who were largely non-smoking (96·6 %) and without children (70·7 %). Most participants bought groceries at other grocery stores (85·7 %), nearly half shopped at farmer's markets (49·8 %) and few shopped at superstores (17·2 %). As purchase quality increased across GPQI tertiles, fewer participants reported shopping at other grocery stores (92·6 % in T1 $v.$ 80·0 % in T3, $P_{\text{trend}} = 0·04$). Unadjusted mean HEI scores increased across GPQI tertiles ($P_{\text{for trend}} < 0·001$) and were descriptively higher for the HEI (HEI$_{BL}$ = 73·2 (9·1) %; HEI$_{MID}$ = 72·4 (9·4) %) compared with the GPQI (54·8 (9·1) %). Experimental group, sex, socio-economic status and household characteristics did not differ across GPQI tertiles. Characteristics in the full study sample (n 224) were compared among those included (n 182) $v.$ excluded (n 42) at 3 months in online Supplementary Table S1. Participants were comparable ($P > 0·05$) in terms of BMI, sex, education, income, smoking, GPQI and HEI$_{BL}$ scores, and experimental group. Participants excluded from the analytic sample were younger (50·9 (13·0)

Table 1. Descriptive characteristics across tertiles of Grocery Purchase Quality Index (GPQI) scores in a sample of shoppers from a single, local grocery store (*n* 209)
(Mean values and standard deviations; percentages)

| | | | GPQI tertile | | | | | | |
| | | | 1 | | 2 | | 3 | | |
Characteristic*	Mean	SD	Mean	SD	Mean	SD	Mean	SD	$P_{\text{for trend}}$†
Age (years)	55·4	13·9	58·1	14·7	54·5	12·7	53·6	14·0	0·05
BMI (kg/m²)‡	25·4	4·6	26·8	5·3	25·4	4·1	24·1	4·0	<0·001
Sex‡									
Female (%)	90·3		92·7		88·4		90·0		0·63
Education‡									
≤Bachelor's (%)	51·0		52·2		47·8		52·9		0·90
Household income‡									
<$100 000 (%)	49·7		60·8		40·4		49·1		0·29
Mean household size	2·3	1·1	2·2	1·1	2·4	1·1	2·3	1·2	0·51
Smoking in household									
Yes (%)	3·4		4·4		2·9		2·9		0·64
Children in household‡									
Yes (%)	29·3		26·9		29·4		31·4		0·56
Shops at other grocery store(s)									
Yes (%)	85·7		92·6		84·3		80·0		0·04
Shops at farmer's market(s)									
Yes (%)	49·8		46·4		48·6		54·3		0·34
Shops at superstore(s)									
Yes (%)	17·2		17·4		22·9		11·4		0·30
Percentage GPQI score	54·8	9·1	45·1	5·2	54·5	2·2	64·8	4·4	<0·001
Percentage baseline HEI score	73·2	9·1	70·5	8·3	71·8	9·3	77·2	8·2	<0·001
Percentage 3-month HEI score‡	72·4	9·4	70·3	9·4	70·5	9·1	76·5	8·5	<0·001
Experimental group									
Intervention (%)	49·8		42·0		57·1		50·0		0·40

HEI, Healthy Eating Index.
* All characteristics were self-reported by primary shopper in the household. HEI, GPQI, experimental group, income, household size, children in household and smoking status reflect household characteristics; age, sex, education, BMI, shopping at farmer's market(s), other grocery store(s) or superstore(s) reflect primary shopper characteristics.
† $P_{\text{for trend}}$ obtained using unadjusted generalised linear models for continuous variables and unadjusted logistic regression for categorical variables.
‡ Missing data were present among BMI (*n* 11), sex (*n* 2), education (*n* 1), household income (*n* 48), children in household (*n* 4), and 3-month HEI score (*n* 27).

v. 56·1 (13·9) years, *P* = 0·03) and had larger household sizes (3·0 (1·5) *v.* 2·2 (1·1), *P* = 0·003).

Table 2 presents age and sex partially adjusted GPQI and HEI correlations at baseline and 3 months, overall and stratified by participant characteristics. At both time points, the GPQI was moderately correlated with the HEI (HEI$_{\text{BL}}$ *r* 0·41, HEI$_{\text{MID}}$, *r* 0·31, *P* < 0·001), indicating acceptable agreement (*r* 0·20–0·49)[27]. Similarly, there was acceptable agreement between the GPQI and HEI at both time points for participants who were female, ≤bachelor's degree and without children. Participants with >bachelor's degree, income ≥$100 000 and normal weight BMI only had acceptable agreement between these measures at baseline. At baseline, good agreement (*r* > 0·49) was observed in participants with children (*r* 0·56) and >65 years old (*r* 0·51). The remaining correlations were not significant at the Bonferroni-adjusted threshold of *P* < 0·0017.

Online Supplementary Table S2 shows the model fit details for each covariate added to the fully adjusted model which had R^2 values of 43·9 % with HEI$_{\text{BL}}$ and 42·0 % with HEI$_{\text{MID}}$, indicating an improvement in variance explained over the unadjusted and age- and sex-adjusted (28·4–30·5 %) models. Fig. 1 displays the multivariable-adjusted concordance between GPQI and HEI scores across GPQI tertiles. In all tertiles and at both time points, GPQI scores were lower than HEI scores and concordance between scores (i.e. mean GPQI score – mean

HEI score) was highest in the healthiest tertile (at baseline, T1 = −26·1 %, T2 = −16·7 %, T3 = 8·2 %; at 3 months, T1 = −26·1 %, T2 = −15·6 %, T1 = −10·0 %; *P* < 0·0001 for both time points).

Table 3 displays multivariable-adjusted mean component score differences at baseline and 3 months across GPQI tertiles. For components except dairy products, concordance increased across tertiles and HEI scores exceeded GPQI scores. In dairy products, concordance decreased across tertiles and HEI scores were lower than GPQI scores (HEI < GPQI by 3·6–2·2 points). Descriptively, concordance was lowest for refined grains (HEI > GPQI by 8·9–7·4 points), dairy products, and greens and beans (HEI > GPQI 3·1–2·0 points), and highest for whole fruit (HEI > GPQI 0·3–1·6 points).

Online Supplementary Tables S3 and S4 compare frequencies and adjusted odds of high (≥80 points) and low (<60 points) HEI scores across GPQI tertiles. High HEI scores attained by 21·1 % of participants at baseline and 18·7 % at 3 months were more frequent among participants in the highest grocery purchase quality tertile (35·6, 41·4 %) *v.* the lowest tertile (8·1, 8·7 %); participants in T3 also had increased odds of HEI scores ≥80 *v.* participants in T1 (HEI$_{\text{BL}}$ OR = 7·32 (2·04, 26·24), HEI$_{\text{MID}}$ OR = 10·55 (1·92, 57·97)). Odds were not significantly different for T2 (HEI$_{\text{BL}}$ OR = 1·84 (0·48, 7·04), HEI$_{\text{MID}}$ OR = 3·07 (0·52, 18·21)). Alternatively, low HEI scores were observed for 8·6 %

Table 2. Partially adjusted (for age and sex) correlations between the Healthy Eating Index-2010 (HEI) and Grocery Purchase Quality Index (GPQI) scores among different characteristics

Characteristic	Baseline			3 months		
	n	GPQI and HEI_{BL} correlation (r)*	P	n	GPQI and HEI_{MID} correlation (r)*	P
Overall	207	0·41	<0·0001†	180	0·31	<0·0001†
Age (years)						
18–45	45	0·46	0·002	37	0·32	0·05
45–65	104	0·28	0·004	89	0·30	0·005
>65	58	0·51	<0·001†	54	0·27	0·05
Sex						
Female	187	0·40	<0·001†	161	0·31	<0·001†
Male	20	0·45	0·05	19	0·34	0·17
Education						
≤Bachelor's	105	0·44	<0·001†	91	0·43	<0·001†
>Bachelor's	101	0·37	<0·001†	88	0·17	0·12
Household income						
<$100 000	80	0·37	0·001†	71	0·32	0·0087
≥$100 000	80	0·46	<0·001†	62	0·25	0·0486
BMI (kg/m^2)						
18·5–25·0 (normal)	99	0·32	0·001†	89	0·14	0·20
25·0–29·9 (overweight)	63	0·32	0·01	53	0·33	0·02
>29·9 (obese)	30	0·46	0·01	27	0·53	0·006
Children in household						
Yes	59	0·56	<0·001†	47	0·36	0·02
No	144	0·35	<0·001†	129	0·30	<0·001†

HEI_{BL}, baseline HEI; HEI_{MID}, 3 months HEI.
* Correlations are partially adjusted for age (except age-stratified correlations) and sex (except sex-stratified correlations).
† P values indicate significance at $P < 0.0017$ (Bonferroni correction for thirty tests at $P < 0.05$).
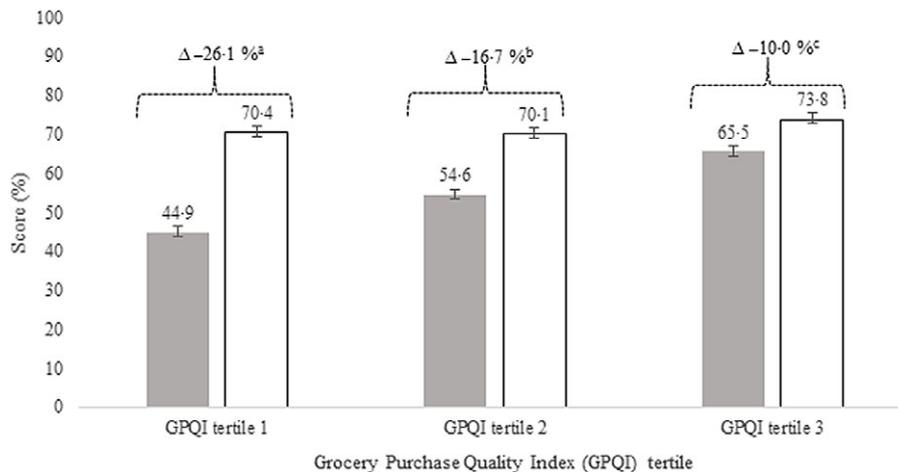
Fig. 1. Score differences* in grocery purchase diet quality (Grocery Purchase Quality Index; GPQI) and individual-level diet quality (Healthy Eating Index; HEI) across GPQI tertiles†. [a,b,c] Unlike letters indicate significant Tukey-adjusted *post hoc* comparisons of score differences across GPQI tertiles; all tertile comparisons of score differences were significant. ■, GPQI; □, HEI. * Score differences (GPQI-HEI, depicted with Δ and brackets) are adjusted for age, sex, BMI and income. For ease of interpretation, GPQI scores and HEI scores (depicted with bars) were derived from separate models adjusted for age, sex, BMI and income. Adjusted HEI scores reflect the HEI_{MID} (3-month) assessment; however, adjusted HEI_{BL} (baseline) measures were similar to HEI_{MID} (tertile 1 = 71, tertile 2 = 71·3, tertile 3 = 75·5). Adjusted differences calculated with the baseline HEI assessment (GPQI − HEI_{BL}) were identical for GPQI tertile 1 (−26·1 %) and similar for tertiles 2 (−15·6 %) and 3 (−8·2 %). † GPQI mean (SE): tertile 1 = 44·9 (0·82); tertile 2 = 54·6 (0·73); tertile 3 = 65·5 (0·75).

of participants at baseline and 8·8 % at midpoint, with fewer in GPQI T3 (2·9, 3·4 %) compared with T2 (7·1, 11·5 %) and T1 (11·3, 15·9 %). There was no difference between T3 and T2 *v.* T1 in odds of having HEI_{BL} or HEI_{MID} scores <60 points in the fully adjusted model.

Weighted κ statistics indicating agreement between tertile rankings of the GPQI and HEI scores were 0·25 at both baseline and 3 months. Both κ statistics obtained from baseline and 3 months were within the fair agreement range ($\kappa = 0.21–0.49$)[27].

Table 4 summarises results and interpretations[27] of the validity measures used to examine agreement between the GPQI and the HEI. Despite weaker correlations at 3 months (r 0·31) compared with baseline (r 0·41), both correlations indicate acceptable agreement, as did both weighted κ coefficients. Score differences for T3 at both time points similarly had

Table 3. Grocery purchase diet quality (Grocery Purchase Quality Index; GPQI) and consumed diet quality (Healthy Eating Index; HEI) component mean scores and mean differences by GPQI tertiles
(Mean values with their standard errors)

Component*	HEI time point	GPQI and HEI correlation (r)		Component score difference (GPQI – HEI)						
				GPQI tertile 1		GPQI tertile 2		GPQI tertile 3		
		r	P	Mean	SE	Mean	SE	Mean	SE	$P_{\text{for trend}}$†
Total fruit	Baseline	0·38	<0·001	−0·7	0·3	−0·5	0·2	0·1	0·2	0·03
	3 months	0·32	<0·001	−0·7	0·3	−0·3	0·3	0·1	0·3	0·07
Whole fruit	Baseline	0·32	<0·001	−1·6	0·3	−1·2	0·2	−0·5	0·2	<0·001
	3 months	0·35	<0·001	−1·4	0·3	−1·0	0·2	−0·3	0·3	0·001
Total vegetables	Baseline	0·20	0·004	−2·1	0·2	−1·2	0·2	−1·1	0·2	<0·001
	3 months	0·18	0·02	−2·0	0·3	−1·1	0·2	−0·8	0·2	<0·001
Greens and beans	Baseline	0·28	<0·001	−2·8	0·3	−2·1	0·3	−2·1	0·3	0·001
	3 months	0·18	0·02	−3·1	0·3	−2·2	0·3	−2·0	0·3	<0·001
Whole grains	Baseline	0·34	<0·001	−2·7	0·6	−1·5	0·6	−0·5	0·6	0·01
	3 months	0·38	<0·001	−3·4	0·6	−1·9	0·5	−1·1	0·6	<0·001
Dairy products	Baseline	0·30	<0·001	2·2	0·7	2·3	0·6	3·0	0·6	0·09
	3 months	0·34	<0·001	2·3	0·7	2·8	0·6	3·6	0·7	0·01
Total protein foods	Baseline	0·34	<0·001	−1·9	0·3	−1·2	0·3	−1·3	0·3	0·04
	3 months	0·20	0·009	−1·5	0·3	−0·9	0·3	−1·1	0·3	0·15
Seafood and nuts	Baseline	0·20	0·009	−2·9	0·4	−2·0	0·3	−1·7	0·3	<0·001
	3 months	0·25	<0·001	−2·6	0·4	−1·7	0·3	−1·9	0·4	0·009
Refined grains	Baseline	0·06	0·40	−8·3	0·5	−8·9	0·4	−7·4	0·4	0·02
	3 months	0·14	0·07	−8·3	0·5	−8·8	0·5	−7·5	0·5	0·21

* HEI seafood/plant protein component includes seafood, nuts, seeds, soya products (other than beverages) and legumes (beans and peas); GPQI seafood/nuts component includes fish, fish products, nuts, nut butters and seeds. Maximum GPQI component scores: total fruit = 5; whole fruit = 5; total vegetables = 5; greens and beans = 5; WG = 10; dairy products = 10; TPF = 5; seafood and nuts = 5; refined grains = 10. Maximum HEI component scores: total fruit = 5; whole fruit = 5; total vegetables = 5; greens and beans = 5; WG = 10; dairy products = 10; TPF = 5; seafood and plant proteins = 5; refined grains = 10.
† $P_{\text{for trend}}$ was obtained using unadjusted generalised linear models.

Table 4. Summary of results from validity measures examined and thresholds indicating agreement

Measure	Interpretation	Value ranges indicating agreement*			Result†
		Poor	Acceptable	Good	
Correlation coefficient	Higher correlations indicate stronger agreement between the measures	<0·20	0·20–0·49	≥0·50	Acceptable
Percentage differences	Smaller differences in percentage scores obtained by the two measures indicate agreement	>10 %	0–10 %		Acceptable (T3), poor (T1 and T2)
Cross-classification‡	Agreement between tertile rankings in each score is desired	<50 % in same tertile		≥50 % in same tertile	Poor (comparing GPQI T3 and high HEI scores, GPQI T1 and low HEI scores)
		>10 % in opposite tertile		≤10 % in opposite tertile	Good (comparing GPQI T1 and high HEI scores, GPQI T3 and low HEI scores)
Weighted κ coefficient	κ Coefficients give a measure of overall relative agreement in tertile rankings	<0·20	0·20–0·60	≥0·61	Acceptable

T3, tertile 3; T1, tertile 1; T2, tertile 2; GPQI, Grocery Purchase Quality Index; HEI, Healthy Eating Index.
* Ranges indicating poor, acceptable and good agreement were adapted from Lombard *et al.*[25].
† Overall results displayed were the same at baseline and 3 months.
‡ In cross-classification, instead of comparing tertile memberships across scores, analyses compared how often participants in GPQI T3 aligned with an *a priori* high (≥80 points) HEI score as well as how often participants with low GPQI scores (T1) aligned with *a priori* low (<60 points) HEI scores.

acceptable agreement, whereas T1 and T2 indicated poor agreement. In cross-classification, <10 % of participants with a high HEI were grouped in the lowest GPQI tertile and <10 % of people with a low HEI were grouped in the highest GPQI tertile, indicating good agreement. Conversely, <50 % of participants in GPQI T3 also had high HEI scores, similarly to participants in GPQI T1 with low HEI scores, indicating poor agreement[27].

Discussion

The present study found that diet quality calculated with partial grocery purchasing data was fairly to moderately concordant with individual-level diet quality, underscoring the promise of using purchasing data as a complementary, objective diet assessment tool[28]. Numerous validation tests (summarised in Table 4) indicate generally acceptable agreement between the GPQI and

HEI. Higher grocery purchase quality was associated with better score concordance and covariate adjustment (age, sex, BMI and income) accounted for >40 % of the variance in models, suggesting that adjustment for covariates may further contribute to the utility of the GPQI as a metric on individual-level diet quality. Interestingly, HEI scores measured at baseline were more strongly predictive of grocery purchase quality than those at 3 months which temporally aligned with the GPQI measurement period.

The acceptable level of agreement between partial purchase data and self-reported dietary intake suggests that combining these measures should be further explored as a means to address measurement error in dietary assessment, which remains a major barrier in nutrition research[29]. Self-reported diet assessment methods are affected by measurement error and bias[30]. Point-of-sale data are an objective measure of household food environments[31,32], and therefore may be useful for correction of measurement error in dietary recalls, similarly to how FFQ can help to account for usual intake variability not captured in multiple 24-h recalls[29]. In this analysis, accounting for socio-demographic variables (age, sex, BMI, income) explained 42–44 % of the total variance in score concordance. Therefore, taking personal characteristics into account likely helped reduce person-specific bias[29].

Score concordance also varied within our sample and was strongest in participants who had the healthiest grocery purchases (i.e. GPQI T3); these participants had significantly higher odds of being identified as having high HEI scores compared to participants with the least healthy grocery purchases (i.e. GPQI T1), but odds of low HEI scores were not different between GPQI tertiles. Participants with the highest concordance and the healthiest grocery purchases reported shopping at other non-study grocery stores significantly less than other participants, indicating that more complete grocery purchasing data or adjustment for the relative completeness may strengthen concordance. Collectively, these results indicate that partial household purchasing data are a viable proxy of individual-level intake, particularly after adjusting for participant characteristics. Further research should explore adjustment for factors related to the completeness of grocery data (e.g. amount of groceries obtained elsewhere) that could improve model fit and concordance.

In the present study, individual-level HEI scores reflecting two different 3-month time frames of self-reported intake were examined. The first assessment took place during study enrollment (HEI$_{BL}$), aligning with the start of the grocery purchasing data collection; the second assessment was taken at the end of the 3-month grocery purchasing data time period (HEI$_{MID}$), aligning the periods of recall and purchase data analysis. Using these two HEI measures, analyses assessed whether HEI$_{BL}$ predicted GPQI scores and whether the HEI$_{MID}$ would prospectively correlate with the GPQI when measurement periods aligned. When assessing validity with two imperfect measures (e.g. GPQI from incomplete purchase data, self-reported HEI), best practices indicate that comparisons should be made within the same time frame[22]. Interestingly, the HEI$_{BL}$ which reflected a time frame not captured in the GPQI was more strongly correlated with grocery purchase quality compared with the HEI$_{MID}$ (r 0·41 $v.$ 0·31). HEI scores at 3 months were 0·8 points lower than baseline, which may be driven by dietary changes in the winter/holiday season[33]; seasonal bias in FFQ can lead to overemphasis of more recent consumption[34]. Therefore, lower agreement at 3 months was likely influenced by seasonal changes in eating and shopping habits around the winter holidays. Additionally, some participants with higher HEI$_{BL}$ and GPQI scores did not complete the HEI$_{MID}$ and were therefore excluded from the 3-month analytic sample which likely contributed to the reduced strength of the overall correlation. Some exploratory analyses (data not shown) suggested that loss of higher scoring participants may explain why some groups had more substantial attenuation of the 3-month correlations (e.g. >bachelor's degree, ≥$100 000 income, normal BMI); among high-income participants, for example, participants included only at baseline scored descriptively higher than participants included at both time points on the HEI$_{BL}$ (73·7 $v.$ 72·0) and GPQI (56·1 $v.$ 54·6).

This analysis contributes to the applications of food purchasing data to complement current methods of diet assessment. Building on previous research by Appelhans et al.[16], who found that HEI scores calculated from complete purchasing data and repeated 24-h recalls were moderately correlated ($\rho c = 0·67$), the present analysis yielded a moderate correlation between partial grocery purchasing data and individual-level diet quality metrics. Present research utilising grocery purchasing data has largely analysed data from third-party companies (e.g. Nielsen) who collect, pool and process data from pre-selected consumer panels[35]. Therefore, our study contributes to the understudied research area of local, non-chain grocery data as a means for assessing diet quality in a nutrition intervention, though consideration must be given to the challenges associated with collecting and analysing grocery purchasing data.

Utilising grocery purchasing data for diet assessment can be analytically challenging due to variation in manufactures' products and point-of-sale systems. Given the promise of grocery purchasing data as a diet quality metric, novel data processing methods should be further explored to address analytic challenges. For example, Tran et al.[36] correctly linked 77–100 % of grocery purchase items to the Food Patterns Equivalent Database groups via utilisation of natural language processing and food concept maps. Similarly, recent efforts have utilised probabilistic and semantic matching to link consumer and retailer data from over 800 000 grocery items to United States Department of Agriculture nutrient composition databases[37]. Further efforts to automate linkages between purchase data and nutritional information are needed for scaling grocery purchase data analysis methods for broader use.

The present study has some limitations. The participating grocery store is an independent specialty market known for its healthier products and has relatively high-income customers who are unlikely to be representative of the general population. Subsequently, the generalisability of our findings is limited by the homogeneity of our sample which had average HEI scores approximately 14 % higher than the general US population[38]. At baseline, participants reported all store types where they purchased groceries during the last month and estimated their total weekly grocery spending across all stores. Participants were not

asked to quantify relative contributions of grocery sources, so models were not adjusted for the relative completeness of the partial grocery purchasing data. However, using estimated weekly grocery spending reported at baseline (data not shown), it can be approximated that 49 % total grocery spending was accounted for at the participating grocery store. Furthermore, while US grocery shoppers visit approximately 4·4 different stores per month, over 90 % of shoppers purchase the majority of their groceries at a single primary store[39], suggesting that the present analysis of grocery purchases from a single store represented a considerable portion of the sample's total grocery purchases. But for 48·9 % of participants who reported shopping at farmer's markets, it is possible that GPQI scores for fruits and vegetables were underestimated. Despite this, >40 % of variance was explained by accounting for demographic characteristics. The present study also had numerous strengths. This analysis comprehensively examined concordance between the GPQI and HEI at two different time points using a variety of validity measures in a relatively large sample. Furthermore, the present study is a major contribution to the diet assessment literature due to our novel examination and comparison of partial grocery data and self-reported dietary data.

In conclusion, this comprehensive analysis found that household-level GPQI scores derived from partial grocery purchasing data were fairly to moderately reflective of the individual-level diet quality of the primary shopper, particularly with adjustment for participant characteristics. Using partial grocery purchase records to validly identify individuals with high and low diet quality may be especially important for dietary interventions by reducing the expenses associated with numerous self-reported dietary assessments and issues-related non-response bias. Future research in more heterogeneous samples is warranted.

Acknowledgements

Supplementary material

For supplementary materials referred to in this article, please visit https://doi.org/10.1017/S0007114520004833

References

1. Afshin A, Sur PJ, Fay KA, *et al.* (2019) Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **393**, 1958–1972.

2. Whelan J, Love P, Millar L, *et al.* (2018) Sustaining obesity prevention in communities: a systematic narrative synthesis review. *Obes Rev* **19**, 839–851.

3. Noy S, Patrick R, Henderson-Wilson C, *et al.* (2018) New frontiers in community initiatives to increase vegetable consumption. *Heal Promot J Aust* **30**, 52–61.

4. Roy R, Kelly B, Rangan A, *et al.* (2015) Food environment interventions to improve the dietary behavior of young adults in tertiary education settings: a systematic literature review. *J Acad Nutr Diet* **115**, 1647–1681.e1.

5. Calancie L, Leeman J, Jilcott Pitts SB, *et al.* (2015) Nutrition-related policy and environmental strategies to prevent obesity in rural communities: a systematic review of the literature, 2002–2013. *Prev Chronic Dis* **12**, 140540.

6. Black C, Moon G & Baird J (2014) Dietary inequalities: what is the evidence for the effect of the neighbourhood food environment? *Health Place* **27**, 229–242.

7. Glanz K & Hoelscher D (2004) Increasing fruit and vegetable intake by changing environments, policy and pricing: restaurant-based research, strategies, and recommendations. *Prev Med (Baltim)* **39**, 88–93.

8. Cade JE (2017) Measuring diet in the 21st century: use of new technologies. *Proc Nutr Soc* **76**, 276–282.

9. Story M, Kaphingst KM, Robinson-O'Brien R, *et al.* (2008) Creating healthy food and eating environments: policy and environmental approaches. *Annu Rev Public Health* **29**, 253–272.

10. McKinnon RA, Reedy J, Morrissette MA, *et al.* (2009) Measures of the food environment. *Am J Prev Med* **36**, S124–S133.

11. Herforth A & Ahmed S (2015) The food environment, its effects on dietary consumption, and potential for measurement within agriculture-nutrition interventions. *Food Secur* **7**, 505–520.

12. Gamba RJ, Schuchter J, Rutt C, *et al.* (2015) Measuring the food environment and its effects on obesity in the United States: a systematic review of methods and results. *J Community Health* **40**, 464–475.

13. Brewster PJ, Guenther PM, Jordan KC, *et al.* (2017) The Grocery Purchase Quality Index-2016: an innovative approach to assessing grocery food purchases. *J Food Compos Anal* **64**, 119–126.

14. 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.

15. Brewster PJ, Durward CM, Hurdle JF, *et al.* (2019) The Grocery Purchase Quality Index-2016 performs similarly to the Healthy Eating Index-2015 in a national survey of household food purchases. *J Acad Nutr Diet* **119**, 45–56.

16. Appelhans BM, French SA, Tangney CC, *et al.* (2017) To what extent do food purchases reflect shoppers' diet quality and nutrient intake? *Int J Behav Nutr Phys Act* **14**, 46.

17. Vadiveloo MK, Guan X, Parker HW, *et al.* (2020) Evaluating the effect of individually-targeted food incentives on grocery purchases: the smart cart study protocol for a randomized controlled cross-over trial. *Contemp Clin Trials* **91**, 105966.

18. Viocare (2020) VioScreen – analyzing dietary behaviors. https://www.viocare.com/ (accessed August 2020).

19. Kristal AR, Kolar AS, Fisher JL, *et al.* (2014) Evaluation of web-based, self-administered, graphical food frequency questionnaire. *J Acad Nutr Diet* **114**, 613–621.

20. Guenther PM, Kirkpatrick SI, Reedy J, *et al.* (2014) The Healthy Eating Index-2010 is a valid and reliable measure of diet quality according to the 2010 Dietary Guidelines for Americans. *J Nutr* **144**, 399–407.

21. Guenther PM, Casavale KO, Reedy J, *et al.* (2013) Update of the Healthy Eating Index: HEI-2010. *J Acad Nutr Diet* **113**, 569–580.

22. Kirkpatrick SI, Baranowski T, Subar AF, *et al.* (2019) Best practices for conducting and interpreting studies to validate self-report dietary assessment methods. *J Acad Nutr Diet* **119**, 1801–1816.

23. Ludbrook J (1998) Multiple comparison procedures updated. *Clin Exp Pharmacol Physiol* **25**, 1032–1037.

24. Hiza HA, Casavale KO, Guenther PM, *et al.* (2013) Diet quality of Americans differs by age, sex, race/ethnicity, income, and education level. *J Acad Nutr Diet* **113**, 297–306.

25. Vanbelle S (2016) A new interpretation of the weighted kappa coefficients. *Psychometrika* **81**, 399–410.

26. Cohen J (1992) A power primer. *Psychol Bull* **112**, 155–159.

27. Lombard MJ, Steyn NP, Charlton KE, *et al.* (2015) Application and interpretation of multiple statistical tests to evaluate validity of dietary intake assessment methods. *Nutr J* **14**, 40.

28. Thompson FE, Kirkpatrick SI, Subar AF, *et al.* (2015) The National Cancer Institute's dietary assessment primer: a resource for diet research. *J Acad Nutr Diet* **115**, 1986–1995.

29. Naska A, Lagiou A & Lagiou P (2017) Dietary assessment methods in epidemiological research: current state of the art and future prospects. *F1000Res* **6**, 926.

30. Kipnis V, Subar AF, Midthune D, *et al.* (2003) Structure of dietary measurement error: results of the OPEN biomarker study. *Am J Epidemiol* **158**, 14–21.

31. Chidambaram V, Brewster PJ, Jordan KC, *et al.* (2013) qDIET: toward an automated, self-sustaining knowledge base to facilitate linking point-of-sale grocery items to nutritional content. *AMIA Annu Symp Proc* **2013**, 224–233.

32. Mamiya H, Moodie EEM & Buckeridge DL (2017) A novel application of point-of-sales grocery transaction data to enhance community nutrition monitoring. *AMIA Annu Symp Proc* **2017**, 1253–1261.

33. Jahns L, Johnson LAK, Scheett AJ, *et al.* (2016) Measures of diet quality across calendar and winter holiday seasons among mid-life women: a 1-year longitudinal study using the automated self-administered 24-hour recall. *J Acad Nutr Diet* **116**, 1961–1969.

34. Subar AF, Frey CM, Harlan LC, *et al.* (1994) Differences in reported food frequency by season of questionnaire administration: the 1987 National Health Interview Survey. *Epidemiology* **5**, 226–233.

35. Bandy L, Adhikari V, Jebb S, *et al.* (2019) The use of commercial food purchase data for public health nutrition research: a systematic review. *PLOS ONE* **14**, e0210192.

36. Tran L-TT, Brewster PJ, Chidambaram V, *et al.* (2017) An innovative method for monitoring food quality and the healthfulness of consumers' grocery purchases. *Nutrients* **9**, 457.

37. Carlson A, Page ET, Zimmerman TP, *et al.* (2019) Linking USDA Nutrition Databases to IRI Household-Based and Store-Based Scanner Data. https://www.ers.usda.gov/publications/pub-details/?pubid=92570 (accessed January 2020).

38. US Department of Agriculture (2018) Average Healthy Eating Index 2015 Scores for Americans by Age Group. https://www.fns.usda.gov/hei-scores-americans (accessed February 2019).

39. Markenson S (2019) How Many Grocery Stores Did You Shop at Last Month? Voices Food Industry Blog. https://www.fmi.org/blog/view/fmi-blog/2019/07/12/how-many-grocery-stores-did-you-shop-at-last-month (accessed October 2020).