Validating the use of photos to measure dietary intake: the method used by DialBetics, a smartphone-based self-management system for diabetes patients

Shigeko Kato¹ · Kayo Waki^{1,2} · Sadako Nakamura^{3,4} · Sanae Osada⁵ · Haruka Kobayashi⁶ · Hideo Fujita¹ · Takashi Kadowaki² · Kazuhiko Ohe⁷

Received: 18 May 2015 / Accepted: 16 September 2015 / Published online: 14 October 2015
© The Japan Diabetes Society 2015

Abstract

Background The accuracy of estimating nutritional intake and balance from photos of meals has not been well documented. However, DialBetics (DB)—our diabetes self-management support system, which is based on information and communication technologies—relies on the photos that type 2 diabetes patients take of their meals with smartphones. Therefore, we designed a study to evaluate this accuracy.

Methods We prepared 61 dishes whose actual amount/value of total energy and each nutrient were known:

protein, fat, carbohydrates, dietary fiber and salt. Their balance—the protein-fat-carbohydrate ratio—was also known, constituting the weighed food record (WFR). Smartphone photos of those dishes were taken, and three registered dietitians evaluated each dish from those photos, naming the dish and estimating the amount of each nutrient in it, plus the dish's balance. These estimated DB and WFR values were compared using the Wilcoxon matched-pairs rank-sum test; intraclass correlation coefficients (ICCs) were calculated. Agreement between the two values for each dish was assessed by Bland-Altman analysis.

Results There were significant ICCs—0.84 for fat (95 % confidence interval 0.75–0.90) and 0.93 for carbohydrates (0.88, 0.96)—but no statistically significant differences

Electronic supplementary material The online version of this article (doi:10.1007/s13340-015-0240-0) contains supplementary material, which is available to authorized users.

✉ Kayo Waki
kwaki-ky@umin.ac.jp
Shigeko Kato
kato-s@umin.ac.jp
Sadako Nakamura
nacsacj@jumonji-u.ac.jp
Sanae Osada
osada@eiyo.ac.jp
Haruka Kobayashi
kobayashiha@nttdocomo.com
Hideo Fujita
fujitah-ky@umin.ac.jp
Takashi Kadowaki
kadowaki-3im@h.u-tokyo.ac.jp
Kazuhiko Ohe
kohe@hcc.h.u-tokyo.ac.jp

- ² Department of Diabetes and Metabolic Diseases, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan
- ³ Institute of Food, Nutrition and Health, Jumonji University, 2-1-28 Sugawara, Niiza-Shi, Saitama 352-8510, Japan
- ⁴ Division of Nutrition Science, Graduate School of Human Health Science, University of Nagasaki, Siebold, 1-1-1 Manabino, Nagayo-cho, Nagasaki 851-2195, Japan
- ⁵ Junior College, Department of Food and Nutrition, Kagawa Education Institute of Nutrition, 3-24-3, Komagome, Toshima-ku, Tokyo 170-8481, Japan
- ⁶ Business Department of Healthcare Business Smart-life Solutions Department, NTT DOCOMO, Inc., 2-11-1 Nagatacho, Chiyoda-ku, Tokyo 100-6111, Japan
- ⁷ Department of Medical Informatics, School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan

¹ Department of Ubiquitous Health Informatics, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan

between DB and WRF for other nutrients or balance. Bland-Altman analysis showed that differences between the two values were random and not biased against nutrient intake; 95 % limits of agreement were acceptable although wide (energy -198 to 210 kcal/dish; carbohydrates -22.7 to 25.8 g/dish).

Conclusion DB's diet evaluation by photos is reliable with apparent potential for assessing diets.

Keywords Type 2 diabetes · Dish photograph · Dietary intake · Telemedicine · Self-management

Abbreviations

DB	DialBetics
WFR	Weighed food record
ICC	Intraclass correlation coefficient
24hR	24-hour recalls
ICT	Information and communication technologies
apps	Smartphone applications
PDA	Personal digital assistant
RD	Registered dietitian
PFC ratio	Protein-fat-carbohydrate ratio

Introduction

Planned diets and self-management are basic necessities if diabetes patients are to maintain glycemic control [1, 2]. Dietitians advise patients about their diet based on each patient's dietary habits as gleaned from weighed food records (WFRs) and 24-hour recalls (24hRs). However, such retrospective methods of ascertaining patients' diets suffer from reliance on each patient's memory and from potential recall bias and misreporting [3]. On the other hand, prospective prescription of set diets can place a high burden on the patient and almost surely means alteration of habitual intake [3].

Information and communication technologies (ICT) make it possible to record patients' exact diet with photos and text data using such devices as mobile phones, smartphones and tablets. Dietary records with digital photos can improve the data quality, consistency and completeness of food recording [4]. Hence, many studies have used meal photos to supplement the more traditional methods of ascertaining food intake for epidemiological studies and clinical dietary surveys of WFRs or 24hRs [4–6]. With the increasingly widespread use of smartphones, various smartphone applications (apps) for dietary records and self-management have been developed [7]. Some apps provide some or all of such functions as monitoring insulin administration, glycemic control, physical activity and recording the diet.

The innovative new methods of assessing dietary intake with mobile devices fall into five categories: electronic food diaries, self-conducted assessments assisted by food photography, 24hR, analysis of food photographs by trained dietitians and automated food photography analysis [5]. A study of an app to facilitate weight loss (My Meal Mate) reported energy intakes that correlated favorably with 24hRs as a means of estimating group intake [3]. Fukuo et al. investigated type 2 diabetes patients and developed an app to let them keep food diaries with photos using a hand-held personal digital assistant (PDA), and the investigators found good agreement between the PDAs and 24hRs in regards to energy and macronutrient intake [8]. These findings suggest that food record apps with photos have good potential as a tool for assessing diets. However, the validity of using smartphone meal photos to measure dietary intake has not been examined with the requisite rigor. Accordingly, whereas previous studies compared estimates of nutritional intake by ICT-based apps with those of (often fallible) memory-based 24hRs, for greater accuracy we compared evaluations by our system with those from WFRs made by trained dietitians.

We developed a smartphone-based self-management support system for type 2 diabetes patients called "DialBetics" (DB) [9]. DB features an automated evaluation and feedback system for each patient's biological data (blood glucose level, body weight and blood pressure), food record and physical activity. We reported that—in a 3-month randomized study—type 2 diabetes patients who used DB significantly improved their HbA1c compared with a control group that did not use DB (-0.4 ± 0.8 %, $p = 0.01$) [10].

In the present study, we investigated the validity of measuring patients' dietary intake by means of photographs of their meals input to DB. The ingredients, weight and nutrient values of the dishes we prepared, photographed by smartphone and input to simulate patients' input, were exactly measured. The known nutrient values of each WFR and the estimated values that came from DB were compared.

Materials and methods

Perhaps the key feature of DB is a dietary evaluation by registered dietitians (RDs) of photos of patient's meals, with those evaluations and some simple messages about dietary habits fed back to the patients. In this study, we conducted a blind test—with photos of dishes we had specially prepared—to determine the validity of evaluating nutritional intake with smartphone photos.

Briefly, unbeknownst to the three RDs who were doing the evaluation, a second set of eight RDs prepared dishes

and took smartphone photos of them; the photos were randomly sent to the evaluating RDs (through DB) as though coming from patients. We could then gauge how well the three RDs estimated the nutritional content from patients' meal photos by seeing how close their estimates were to the known values of the RD-prepared dishes. It should be re-emphasized that the three evaluating RDs did not know that this test was being conducted.

Selection criteria and classification of dishes

The eight RDs prepared 61 dishes under the direction of the study manager. The dishes included ones that had actually been input to DB by type 2 diabetes patients and others that are popular dishes commonly available in Japan from the food service and catering industries and in regular and fast-food restaurants. Popular cooking methods were used, Japanese, Chinese and Western cooking styles, plus those of other international cuisines and a mixture of styles. The dishes used in the study are shown in Supplement 1. There were four groups of dishes, all conforming to the Japanese Food Guide Spinning Top: grain dishes ($n = 26$), main dishes ($n = 16$), side dishes ($n = 13$) and soups ($n = 6$) [11]. The classification of these groups was based on each dish's main ingredients. There were three grain dish groups: rice dishes (Supplement 1, nos. 1–5), noodles (nos. 6–8) and wheat dishes (no. 9). Main dishes were classified by the cooking method and main ingredients: meat, fish, eggs and soybeans (nos. 10–16). Side dishes were classified the same way: vegetables, seaweed and mushrooms (nos. 17–21). Soups (no. 21) were classified by their kind of stock and ingredients. On average, 59 of the specially prepared dishes included 5.4 different foods/ingredients (excluding seasonings, spices, alcohol and oil for cooking); the other four dishes comprised a single food: brown rice or boiled soybeans.

Weighed food record for dishes

There was a WFR for each of the prepared dishes. All ingredients, including seasonings, were measured with a digital cooking scale (Shimadzu PZ-2000, Shimadzu Corp., Kyoto, Japan), measuring cup or measuring spoon. The RDs recorded the name of each dish and its ingredients, with measured weights. The energy and nutrient contents of each dish were calculated this way: all ingredients were given the food code found in the standard tables of food composition for Japan, 2010 edition; based on that code, nutrient values for the measured weights were analyzed by Excel-Eiyokun software (version 6.0, Kenpakusha Corp., Tokyo, Japan). The sum of nutrient values was recorded for each dish. We thus knew the exact energy content and amount/values of five nutrients—protein, fat, carbohydrate,

dietary fiber and salt—and their balance, i.e., the protein-fat-carbohydrate (PFC) ratio of each dish, data commonly requested in diet studies and the focus in the prevention of hypertension.

Photograph preparation

The prepared dishes were laid out in the meal combinations and patterns usual in Japan (with chopsticks or cutlery for authenticity) on a tablecloth with a typical 4.5×4.5 -cm square pattern. The photos to be input to DB were taken by smartphone at a 45 – 60° angle to show the depth of dishes, making sure that each dish was fully visible. Each dish was used only one time, and dishes were grouped to constitute typical meals. In all, 24 meal photos were taken: meals composed of two dishes ($n = 13$), three dishes ($n = 9$) and four dishes ($n = 2$).

Dietary evaluation using DialBetics

From the photos input to DB, three trained RDs identified each dish in the meal and estimated its ingredients and the method used to cook it. They also estimated each portion size, then combined all those data to evaluate dietary intake. They calculated nutrient values using the DB database. Once they had identified each dish from the meal photo, they chose those dishes (or their equivalent—i.e., similar ingredients and cooking methods) from the database, then calculated the nutrient values for the portion size according to the database. As noted, these RDs had not prepared the dishes and did not know about their origin or even that this trial was being conducted.

Because the three RDs had been specifically trained to evaluate dietary intake using meal photos—and had experience in doing so for various businesses, with their estimates known to fall within acceptable variation ranges—the photos (both from the patients and for the blind study) could be submitted randomly, each photo going to only one RD for evaluation. Because of their training and proven experience there was no need for multiple evaluations of photos.

Identifying each dish by name

Obviously, for accurate evaluation of dietary intake, RDs must be able to identify, from photos, the name of each dish in a meal and estimate its nutrient values. This identification must include accurate identification of the dish components: the ingredients and cooking method [12]. Therefore, we defined accurate identification of photo meals as being only those that matched the meals listed in Supplement 1 or were equivalent to those meals in the main ingredients and cooking method used to prepare them. For

example, if the dietitian identified a dish as “vegetable tempura,” but in fact it was “shitake mushroom tempura,” we counted that as accurate. If the dietitian used a name different from that in the database, but it was clearly the same dish, this was also accurate. Any other identification was deemed inaccurate.

Statistical analysis

Analyses were performed using STATA statistical software, release 12 (Stata Corp.). The Wilcoxon matched-pairs rank-sum test was used to compare group means of energy (kcal) and macronutrient value as measured by both WFR and DB.

Correlation between the two methods for determining each energy and macronutrient value was assessed by intraclass correlation coefficient (ICC) (one-way analysis of variance, absolute agreement and individual measures). For ICC analyses, all data were log-transformed to improve the skewed distribution. Agreement between the two methods was assessed using Bland-Altman regression; the mean differences in the energy and macronutrient value between the methods were plotted against the average value, and the 95 % limits of agreement were marked.

Results

Accuracy of DialBetics compared with the weighed food record

Table 1 shows the median [95 % confidence interval (CI)] (untransformed data) intake of energy (as kcal), protein (in grams), fat (g), carbohydrates (g), dietary fiber (g), salt

(g) and PFC ratio calculated for each nutrient as recorded by DB and WFR for the same dishes. The ICCs were statistically significant for energy and all nutrients (transformed data). Energy recorded by DB correlated well with that of the WFR (energy: ICC 0.92, 95 % CI 0.87–0.95, $p < 0.001$); correlations for macronutrients were high and statistically significant, 0.84 (0.75–0.90) for fat and 0.93 (0.88–0.96) for carbohydrates ($p < 0.001$). When all dishes were analyzed, there were no statistically significant differences between DB and the WFR for any of the nutrients or the PFC ratio. DB recorded higher macronutrient values than the WFR except for energy and protein.

Agreement between the two methods

Figure 1 shows the Bland-Altman plot for the agreement between DB and the WFR; the mean differences are small but the 95 % limits of agreement are wide for energy and macronutrients (energy −198 to 210 kcal, protein −10.7 to 9.8 g, fat −13.0 to 13.7 g, carbohydrate −22.7 to 25.8 g, dietary fiber −2.2 to 2.6 g, salt −2.8 to 2.8 g). Although plots for each nutrient show a few outliers, the majority of the measurements are scattered along the equality line. Thus, the plots show good agreement between the two methods. They indicate that the differences were random and did not demonstrate any systematic bias.

Effects of accurate dish evaluation on the estimation of nutrient intake

Table 2 summarizes the mean total energy and macronutrients recorded by DB and the WFR for correctly evaluated dishes. It also shows the PFC ratios

Table 1 Comparison of dietary intakes per dish between DialBetics and the WFR for all samples ($n = 61$)

	DialBetics	WFR	DialBetics–WFR	<i>P</i> value	ICC	<i>P</i> value for ICC
	Median (P5, P95)				<i>r</i> (95 % CI)	
Energy (kcal)	268 (25, 647)	266 (40, 600)	−7 (−148, 203)	0.95	0.92 (0.87, 0.95)	<0.001
Protein (g)	10.0 (1.4, 29.0)	10.8 (1.6, 27.5)	−0.1 (−10.6, 6.9)	0.77	0.90 (0.84, 0.94)	<0.001
Fat (g)	9.6 (0.4, 23.7)	7.3 (0.3, 23.7)	0.2 (−9.6, 14.1)	0.47	0.84 (0.75, 0.90)	<0.001
Carbohydrate (g)	26.9 (3.2, 91.4)	23.9 (1.9, 79.0)	0.7 (−18.0, 18.2)	0.38	0.93 (0.88, 0.96)	<0.001
Dietary fiber ^a (g)	1.9 (0.5, 6.1)	2.0 (0.1, 4.7)	0.2 (−1.5, 2.3)	0.22	0.73 (0.59, 0.83)	<0.001
Salt ^a (g)	1.9 (0.4, 5.3)	1.7 (0.4, 5.2)	0.0 (−2.8, 2.4)	0.75	0.48 (0.26, 0.65)	<0.001
Protein (% E)	15.9 (8.8, 40.0)	17.3 (7.0, 36.1)	−0.5 (−13.2, 8.8)	0.32		
Fat (% E)	30.8 (5.6, 59.7)	28.8 (5.4, 64.2)	0.8 (−16.3, 18.7)	0.58		
Carbohydrate (% E)	51.2 (15.0, 86.2)	49.8 (11.5, 79.8)	0.1 (−16.0, 22.2)	0.52		

Compared between DialBetics and the WFR assessed by Wilcoxon matched-pairs rank-sum test. ICC analyses were conducted on transformed (natural log) data

WFR weighed food record, ICC intraclass correlation coefficient, P5 5th percentile, P95 95th percentile, % E % energy

^a Dietary fiber and salt values were analyzed by 60 samples

Fig. 1 Bland-Altman plot for nutrient values calculated from DialBetics (DB) and the weighed food record (WFR). **a** Plot for energy, the limits of agreement are -198 to 210 kcal, $p = 0.30$; **b** plot for protein, the limits of agreement are -10.7 to 9.8 g, $p = 0.50$; **c** plot for fat, the limits of agreement are -13.0 to 13.7 g, $p = 0.67$; **d** plot for carbohydrates, the limits of agreement are -22.7 to 25.8 g, $p = 0.22$; **e** plot for dietary fiber, the limits of agreement are -2.2 to 2.6 g, $p = 0.15$; **f** plot for salt, the limits of agreement are -2.8 to 2.8 g, $p = 0.19$. All plots showed no systematic errors

calculated for each method. It also shows all the same data for the incorrectly evaluated dishes. When dishes were analyzed separately, depending on whether they were correctly or incorrectly evaluated, the energy and macronutrient values recorded by DB for correctly evaluated dishes correlated well with those of the WFR (energy ICC 0.92, 95 % CI 0.87–0.95, $p < 0.001$; protein ICC 0.87, 95 % CI 0.78–0.92, $p < 0.001$; fat ICC 0.74, 95 % CI 0.59–0.84, $p < 0.001$; carbohydrate ICC 0.94, 95 % CI 0.90–0.96, $p < 0.001$). However, for incorrectly evaluated dishes, although energy, protein and carbohydrates were significantly correlated with those measured by the WFR, they were fairly wide with 95 % CI (energy ICC 0.73, 95 % CI 0.17–0.94, $p = 0.009$; protein ICC 0.91, 95 % CI 0.64–0.98, $p < 0.001$; carbohydrate ICC 0.78, 95 % CI 0.28–0.95, $p = 0.004$). There were no significant differences in the daily energy intake, macronutrients or PFC ratio between the correctly and incorrectly evaluated dishes.

Discussion

The accuracy of estimating the nutritional intake and balance from photos of meals has not been well studied. Therefore, it was noteworthy that, in the present study, with the same dishes, the correlations between DB and the WFR for energy, macronutrients, dietary fiber and salt were moderate to high with ICCs from 0.84 to 0.93 and a narrow range of 95 % CIs. With individual dishes, Bland-Altman plots showed fairly wide limits of agreement for total energy and other nutrients. On the group level, when dishes were analyzed all together by calculating the mean of total energy and each nutrient of all dishes, total energy and other nutrients were not statistically significantly different. Thus, DB seems to have potential as a dietary assessment tool.

To our knowledge, only one previous study has assessed the validity of data about nutrient intake collected by a smartphone application and compared with a reference

Table 2 Comparison of dietary intake per dish between DialBetics and the WFR by accuracy of food image evaluation

	DialBetics	WFR	DialBetics–WFR	<i>P</i> value	ICC	<i>P</i> value for ICC
	Median (P5, P95)				(95 % CI)	
Correctly evaluated dishes (<i>n</i> = 53)						
Energy (kcal)	268 (25, 675)	266 (31, 634)	−1 (−135, 149)	0.98	0.92 (0.87, 0.95)	<0.001
Protein (g)	9.5 (1.7, 29.2)	10.8 (1.9, 27.5)	−0.1 (−10.6, 6.9)	0.86	0.87 (0.78, 0.92)	<0.001
Fat (g)	9.6 (0.4, 23.7)	8.3 (0.2, 29.6)	0.2 (−9.9, 7.7)	0.48	0.74 (0.59, 0.84)	<0.001
Carbohydrate (g)	23.9 (2.6, 94.4)	23.9 (1.6, 81.8)	1.3 (−18.0, 18.2)	0.26	0.94 (0.90, 0.96)	<0.001
Dietary fiber ^a (g)	1.8 (0.4, 6.3)	2.0 (0.1, 4.7)	0.2 (−1.6, 2.2)	0.34	0.73 (0.58, 0.84)	<0.001
Salt ^a (g)	2.0 (0.1, 5.4)	1.8 (0.4, 5.5)	0.0 (−3.0, 2.6)	0.64	0.73 (0.58, 0.84)	<0.001
Protein (% E)	15.9 (6.9, 43.8)	16.0 (6.8, 36.2)	−0.6 (−13.2, 8.8)	0.23		
Fat (% E)	28.8 (5.5, 60.5)	26.6 (5.2, 69.0)	0.8 (−16.3, 18.7)	0.73		
Carbohydrate (% E)	49.9 (14.0, 86.2)	50.1 (10.8, 80.0)	0.2 (−16.1, 20.3)	0.55		
Incorrectly evaluated dishes (<i>n</i> = 8)						
Energy (kcal)	261 (14, 578)	253 (55, 483)	−36 (−257, 254)	1.00	0.73 (0.17, 0.94)	0.009
Protein (g)	10.8 (1.2, 25.2)	13.6 (1.0, 29.3)	0.0 (−10.9, 9.7)	0.73	0.91 (0.64, 0.98)	<0.001
Fat (g)	9.5 (0.1, 26.7)	5.2 (2.9, 18.9)	1.7 (−7.7, 14.6)	0.58	0.42 (−0.30, 0.85)	0.117
Carbohydrate (g)	32.5 (3.2, 75.4)	29.3 (5.1, 59.7)	−4.8 (−19.2, 38.4)	0.89	0.78 (0.28, 0.95)	0.004
Dietary fiber (g)	2.2 (1.1, 5.9)	1.8 (0.0, 4.3)	0.6 (−1.3, 2.5)	0.26	0.48 (−0.23, 0.87)	0.082
Salt (g)	1.3 (0.4, 3.1)	1.6 (0.4, 2.6)	−0.3 (−0.7, 1.3)	0.67	0.80 (0.34, 0.96)	0.003
Protein (% E)	16.1 (10.7, 40.0)	18.1 (7.0, 31.0)	0.5 (−13.7, 32.7)	0.78		
Fat (% E)	33.0 (6.4, 48.7)	31.1 (13.4, 58.9)	3.4 (−52.5, 19.4)	0.67		
Carbohydrate (% E)	56.2 (25.8, 91.4)	49.0 (32.3, 64.8)	−4.4 (−11.5, 54.3)	0.67		

Compared between DialBetics and the WFR assessed by Wilcoxon matched-pairs rank-sum test. ICC analyses were conducted on the transformed (natural log) data

WFR weighed food record, ICC intraclass correlation coefficient, P5 5th percentile, P95 95th percentile, % E % energy

^a Dietary fiber and salt values were analyzed by 52 samples

measure. The ICCs of our study compare favorably with those of that previous study. Carter et al. compared My Meal Mate (MMM)—a smartphone application designed to support weight loss by permitting self-monitoring of food and drink intake—to a reference measure of 24hRs. Fifty volunteers recorded their food and drink intake with MMM for 7 days [3]. Daily totals for energy and macronutrients recorded by MMM were compared against the corresponding day of recall using *t* tests for group means and Pearson's correlations. Bland-Altman analysis was used to assess the agreement between the two methods. MMM correlated well with 24hRs for estimating group means with Pearson's correlations (*r*) of 0.77 (95 % CI 0.62–0.86) for day 1 and *r* 0.85 (95 % CI 0.74–0.91) for day 2. However, there was wide disagreement between what MMM and 24hR found as daily energy totals: the limit of agreement was −3378 to 3243 kJ/day (−807 to 775 kcal/day) on day 1 for total energy. This anomalous discrepancy in findings can be best explained by the fact that the largest outlying differences recorded by the two methods when examining individuals' intake were clearly attributable to mistakes in estimating portion sizes.

The same would hold true for our study. However, our study shows that correct identification of the name of a food or dish tends to lead to correct estimation of portion size, which, in turn, results in more precise estimation of energy and macronutrients for each dish.

A few studies that compared the accuracy of intake measurement by electronic devices with a reference measure reported agreement as analyzed by Bland-Altman analysis in addition to the mean of the 61 dishes. Validation of DietMatePro—an integrated Web and PDA app—involved a Bland-Altman plot of the agreement between the app and 24hR in a study with 28 adults. The limits of agreement as to daily energy totals were fairly wide, showing −1600 to 1600 kcal with low to moderate correlation of the coefficient of *r* 0.520–0.797 [13]. For validation of food photography, 24hR had high *r* of 0.99 for daily energy totals with acceptable 95 % limits of agreement of −683 to 206 kcal/day [14]. The limits of agreement for the mean of 2 days of MMM were −582 to 483 kcal with moderate *r* of 0.85 (95 % CI 0.74–0.92) [3]. Since all those studies assessed daily total energy intake—whereas the present study assessed the total energy intake

of one dish—the difference in limits of agreement between their energy intake totals and ours is understandable: such limits are naturally higher for daily total intake than for the energy totals of individual dishes. However, given the relatively higher correlations of total energy for the present study compared with those of the previous studies, similarly wide but acceptable limits of agreement could be expected for the present study.

WFRs were used as a reference in the present study in an attempt to make correlations more exact. They provided a precise referent against which to measure the accuracy of estimates based on processing of the photos that were input to DB and the digital food photography records in the DB database. The most usual Japanese meals of breakfast, lunch and dinner consist of a combination of a few elements—a rice bowl, soup, main dish and side dish. Accurate identification of each dish's name and portion size is obviously required in order to measure the food and nutrient intake for individuals. Notwithstanding the high ICC between DB and the WFR for all dishes, ICCs were of course much higher and significant for correctly evaluated dishes than for incorrectly evaluated dishes. Conversely—and unsurprisingly—ICCs in energy and each nutrient between DB and the WFR were lower for incorrectly evaluated dishes than for correctly evaluated dishes with a wide range. This may be due to the variance that stems from the high variability among incorrectly identified dishes (i.e., the portion size of each dish was more likely to be incorrectly measured when the name of a dish was incorrectly identified)—a mistake subsequently reflected in high differences in estimates of energy and nutrient intake. Thus, correct identification of dishes is a *sine qua non* for accurate diet evaluation with DB.

The strengths of the present study include comparison of DB to the WFR, evaluation of a common Japanese food pattern and a test of the ability of DB to accurately measure nutrient intake in the normal environment expected for the app usage. A potential limitation is the relatively small number of dishes used for the study and the fact that they were mostly Japanese dishes, limiting generalizability to other populations. In addition, use of DB—especially its food recording function—declined over many weeks, limiting the number of dishes for evaluation [10]. Validity of DB could be affected if the number of dishes recorded was too small to conduct group-level dietary assessment.

As a dietary evaluation tool, BD might be viewed as expensive because dietitians conduct the evaluation. Automating the dietary assessment module will cut costs while making feedback to patients about diet and suggested modifications timelier. Our work continues to investigate the reduction of cost and time by automating DB's dietary assessment module.

In conclusion, the DB dietary evaluation function is associated—and in agreement—with the WFR, validating its use of meal photos to evaluate patient nutritional intake, at least in Japan. Further testing with other (and larger) study populations using different food patterns is needed to understand the broader clinical applicability of DB.

Acknowledgments This work was funded by NTT DOCOMO and was partly supported by JSPS KAKENHI grant nos. 23790559 and 24700843. The authors thank Dr. Tsuneyuki Oku and Dr. Katsuhisa Omagari, University of Nagasaki, for their valuable advice.

Compliance with ethical standards

Conflict of interest SK is a member of the Department of Ubiquitous Health Informatics, which is engaged in a cooperative program between the University of Tokyo and NTT DOCOMO. KW is a member of the Department of Ubiquitous Health Informatics, which is engaged in a cooperative program between the University of Tokyo and NTT DOCOMO. HF is a member of the Department of Ubiquitous Health Informatics, which is engaged in a cooperative program between the University of Tokyo and NTT DOCOMO. HK is a member of NTT DOCOMO. SN, SO, TK and KO declare that they have no conflict of interest.

Ethical standards This article does not contain any studies with human or animal subjects performed by the any of the authors.

References

1. Pastors JG, Warshaw H, Daly A, Franz M, Kulkarni K. The evidence for the effectiveness of medical nutrition therapy in diabetes management. *Diabetes Care*. 2002;25:608–13.
2. UKPDS Group. UK Prospective Diabetes Study 7: response of fasting plasma glucose to diet therapy in newly presenting type II diabetic patients. *Metabolism*. 1990;39:905–12.
3. Carter MC, Burley VJ, Nykjaer C, Cade JE. 'My Meal Mate' (MMM): validation of the diet measures captured on a smartphone application to facilitate weight loss. *Br J Nutr*. 2013;109:539–46.
4. Ngo J, Engelen A, Molag M, Roesle J, Garcia-Segovia P, Serra-Majem L. A review of the use of information and communication technologies for dietary assessment. *Br J Nutr*. 2009;101(Suppl 2):S102–12.
5. Sharp DB, Allman-Farinelli M. Feasibility and validity of mobile phones to assess dietary intakes. *Nutrition*. 2014;30:1257–66.
6. Imai T, Otsuka R, Katou Y, Nakamura M, Ando F, Shimokata H. Advantages of taking photographs with the 3-day dietary record. *J Integr Stud Diet Habits*. 2009;20:203–10.
7. El-Gayar O, Timsina P, Nawar N, Eid W. Mobile applications for diabetes self-management: status and potential. *J Diabetes Sci Technol*. 2013;7:247–62.
8. Fukuo W, Yoshiuchi K, Ohashi K, Togashi H, Sekine R, Kikuchi H, Sakamoto N, Inada S, Sato F, Kadowaki T, Akabayashi A. Development of a hand-held personal digital assistant-based food diary with food photographs for Japanese subjects. *J Am Diet Assoc*. 2009;109:1232–6.
9. Waki K, Fujita H, Uchimura Y, Aramaki E, Omae K, Kadowaki T, Ohe K. DialBetics: smartphone-based self-management for type 2 diabetes patients. *J Diabetes Sci Technol*. 2012;6:983–5.

10. Waki K, Fujita H, Uchimura Y, Omae K, Aramaki E, Kato S, Lee H, Kobayashi H, Kadowaki T, Ohe K. DialBetics: a novel smartphone-based self-management support system for type 2 diabetes patients. *J Diabetes Sci Technol*. 2014;8:209–15.
11. Yoshiike N, Hayashi F, Takemi Y, Mizoguchi K, Seino F. A new food guide in Japan: the Japanese food guide spinning top. *Nutr Rev*. 2007;65:149–54.
12. Takachi R, Kudo Y, Watanabe S, Kimira M. Validity of smaller sample sizes for the dish-based component table to use with the self-reported “dietary record by cooked dishes”. *Jpn J Nutr Diet*. 2006;64:97–105.
13. Beasley J, Riley WT, Jean-Mary J. Accuracy of a PDA-based dietary assessment program. *Nutrition*. 2005;21:672–7.
14. Lazarte CE, Encinas ME, Alegre C, Granfeldt Y. Validation of digital photographs, as a tool in 24-h recall, for the improvement of dietary assessment among rural populations in developing countries. *Nutr J*. 2012;11:61.