Validation of an automated wearable camera-based image-assisted recall method and the 24-hour recall method for assessing maternal and child dietary diversity and women's time use in rural Uganda

Andrea L. Spray Bulungu

Thesis submitted in accordance with the requirements for the degree of
Doctor of Philosophy of the University of London

2022

Department of Population Health
Faculty of Epidemiology and Population Health
London School of Hygiene and Tropical Medicine, University of London

Funding Details: This project was funded by a competitive research grant through the Innovative Methods and Metrics for Agriculture and Nutrition Actions (IMMANA) programme. IMMANA is co-funded by UK Foreign Commonwealth and Development Office (FCDO) and by the Bill & Melinda Gates Foundation. IMMANA Grants are led by the London School of Tropical Hygiene and Medicine (LSHTM).

Declaration of own work

I, Andrea L. Spray Bulungu, confirm that the work presented in this thesis is my own. Where information has been gathered from other sources, I confirm that this has been indicated in the thesis.

Signed: [redacted] Date: 26 June 2022

Abstract

Retrospective methods for assessing maternal and child dietary diversity and women's time-use are labour-intensive and prone to misreporting. This thesis comprises the first-ever validation of an automated wearable camera (AWC), an inexpensive technology for prospectively and unobtrusively collecting data, in a low-income country (LIC). The cross-sectional study was conducted in rural Uganda with mothers of a child aged 12-23 months (n=211) to (1) assess the feasibility of an AWC-based image-assisted recall (AWC-IAR); and (2) evaluated the concurrent criterion validity of the AWC-IAR method and the 24-hour recall (24HR) method, compared to observation for (a) assessing maternal and child dietary diversity, and (b) women's time use. Agreement was assessed using the Bland–Altman limits of agreement (LOA) approach.

Most participants rated the AWC-IAR method as good or very good. Eight participants withdrew from the study and 27 were excluded due to AWC inoperability. The relative bias was low, for maternal and child dietary diversity estimated using semi-quantitative multiple-pass 24HR and image-assisted recall (IAR) methods, but the LOA were wide. For time use, the systematic bias varied from 1 minute (domestic chores) to 226 minutes (caregiving) for 24HR and 1 minute (own production) to 109 minutes (socializing) for IAR. The LOA were within 2 hours for employment, own production, and self-care for 24HR and IAR but exceeded 11 hours (24HR) and 9 hours (IAR) for caregiving and socializing.

The AWC-IAR method was feasible in this rural LIC context, although challenges remain. The IAR and 24HR methods provided an accurate estimate of median maternal and child dietary diversity and minutes women spend in employment, own production, and domestic chores. The median time women spent caregiving and socializing was underestimated by both methods. The high LOA indicate that, for research purposes, associations between either dietary diversity or time use and outcomes would be attenuated.

Acknowledgements

Thank you everyone who supported me at LSHTM, most especially my supervisors, **Elaine Ferguson** and **Luigi Palla** (now at the University of Roma La Sapienza). Elaine, I am so grateful, especially, for your kindness, and so inspired by your passion and dogged commitment to high quality research. Luigi, I am thankful for your statistical guidance, of course, but also for your encouragement during a few periods when my stamina flagged. Thank you, too, to my supervisor, **Suneetha Kadiyala**, and to my advisor, **Sujit Rathod**, who both provided timely and practical guidance along the way. Thank you, **Jenny Fleming** and **Lauren Dalton,** for always cheerfully responding to every (and there were many) administrative question I had.

I also wish to thank the other members of the research team, including co-PIs **Kate Wellard** and **Joweria Nambooze**, and **Jan Priebe**, **Lora Forsythe**, **Pamela Katic**, and **Gwen Varley**. Thank you too to **Stephen Young** at NRI, who provided statistical guidance early on, and to the **Sasakawa** team who provided entree into the communities of Bugiri and Kamuli. I hold a special place in my heart for **Bernice Galinda** and **Nakimuli Sarah**, who's giggling (sometimes with me and sometimes at me) got me through months of seemingly endless data entry. I am grateful, too, to the women who participated in this study, and their families and communities.

Long before I embarked on this PhD journey, well-timed and caring mentorship helped me build the skills and confidence I would need to develop into the researcher I've become. I owe each of you a hearty debt of gratitude. **Mr. Adams** volunteered me into a team (of all boys, if I remember correctly) building a weight-bearing tower out of balsa wood and marshmallows. I didn't get it then, but I get it now. **Mr. Ikenn** introduced me to Pink Floyd as well as pre-calculus, making the latter, at least, more relatable than it otherwise would have been. **"Dr" Rule** nurtured me through not one but two levels of A.P. physics, while I was simultaneously wracked by the fear of failure yet trying desperately to maintain my "artist's" air of cool. **Joe Trickey** goaded me into taking an HTML class, before promptly dropping out himself, which amounted to my first exposure to computer programming. As an art student warily dipping a toe into computer science, **Suzanne Menzel** made sure to give me versions of in-class quizzes with concrete - rather than abstract - word problems, knowing long before I did that if I could relate to a problem, I could solve it. I went on to get a CS minor. **Robert Pless**, without whom

I never would have had the gumption to pursue a graduate degree, showed me what a life dedicated to research looked like. **Julie Jacobs** introduced me to the world of public health and, so doing, helped me see how my web design and usability skills, honed in the corporate world, could be used for good. It is no exaggeration to say that she changed my life. At Washington University in St Louis, where I completed an MPH, **Ross Brownson** gave me hope; **Matt Kreuter** made it impossible NOT to take the leap; and **Lora Iannotti** gave me wings. In DC, **Leslie Elder** provided me a safe and secure environment to plant some roots and grow, as a human and a nutritionist. Lora and Leslie never let me forget that they believed PhD was the path for me, while (somewhat) patiently waiting for me to reach the same conclusion.

Sometimes growth comes from adversity, and so I am also grateful to the people who helped me build strength in other ways. The professor who told me, an undergrad struggling to find the bridge between art and computer programming, that I didn't need a computer science degree to use Photoshop. The teaching assistants who chuckled smugly when I, a first-year grad student, spoke of my aspiration to work in international public health. "The fifty". I would have quit the PhD before I began were it not for **Todd Grizzle**, who convinced me that a PhD does not a bag lady make, and a community of women who walked this path before me, and were always there with advice, when I needed it most. **Anne Hamtil Roux**, **Julie Ruel-Bergeron**, **Jody Harris**, **Alissa Pries, Anna Herforth -** thank you. And I hazard to think of where I would be without the support of **Nancy Williger**, **Elaine Klionsky** and **Sadia Cowen**.

Most of all, I am grateful for the love and support of my family and chosen family, most especially my mom (**Anne Jerands**) and dad (**Joel Spray**), and my husband (**Bulungu Isaac**), **Julie Hutchins**, **Shae Wyatt**, **Li Zhou**, **Timothy** and **Margaret Hower**, **Manisha Tharaney**, **Aaron Buchsbaum**, and **Monika Gutestam**. I love you. I also thank the many family members, friends, acquaintances, colleagues, and perfect strangers who have cheered me on all along the way. And finally, I am grateful for Uganda, which provided fertile ground for this work and where I've finally found a home. Akwata mpola atukka wala. It truly takes a global village to raise a Doctor of Philosophy.

To those who believed in me, and those who did not.

"You can't remember something that never made an impression."

- Watts, Reminiscence (2021)

Table of Contents

Abbreviations

24h	24-hour
24HR	24-hour recall
AfrII	Africa Innovations Institute
AI	artificial intelligence
AWC	automated wearable camera
AWC-IAR	AWC-based image-assisted recall
BAPs	Bland-Altman Plots
BMGF	Bill & Melinda Gates Foundation
BMI	body mass index
CAPI	Computer-Assisted Personal Interview
DDS	dietary diversity score
DHS	Demographic and Health Survey
FAO	Food and Agriculture Organization of the United Nations
GDP	gross domestic product
GPS	global positioning system
HETUS	Harmonised European Time Use Surveys
HIC	high-income country
IAR	image-assisted recall
ICATUS	International Classification of Activities for Time-Use Statistics
ICTs	Information and Communication Technologies
ID	identification number
IM	Innovative Methods questionnaire
IMMANA	Innovative Methods and Metrics for Agriculture and Nutrition Actions
IRB	institutional review board
IVR	interactive voice response
LIC	low-income country
LMIC	lower-middle-income country
LOA	limits of agreement
LSHTM	London School of Hygiene and Tropical Medicine
LSMS	Living Standards Measurement Study
MDD	Minimum Dietary Diversity
MDD-W	Minimum Dietary Diversity for Women of Reproductive Age
NRI	University of Greenwich Natural Resources Institute
OBS	direct observation
PIs	principal investigators
PPI	Poverty Probability Index
RA	research assistant
SAA	Sasakawa Africa Association
SD	standard deviation
TUD	time-use diary
UBOS	Uganda Bureau of Statistics

UK	United Kingdom
UMIC	upper-middle-income country
UNCST	Uganda National Council for Science and Technology
UNICEF	United Nations Children's Fund
USA	United States of America
WEAI	Women's Empowerment in Agriculture Index
WFR	weighed food record
WHO	World Health Organization

List of Tables

List of Figures

Chapter 1: Introduction

Background

Undernutrition poses severe long-term consequences for women's and children's development, health, and well-being globally. Nearly one-half (45%) of all child deaths are attributable to undernutrition. (1) Suboptimal diet is among the leading causes of death globally. (2,3) The global economic cost of all forms of malnutrition is estimated to be approximately US$3.5 trillion per year, and an estimated annual gross domestic product (GDP) loss of up to 12% occurs in poor countries due to poor nutrition. (4,5)

Globally, the prevalence of child stunting is decreasing, however 22% (149 million) of children under 5 years of age remain stunted. (6) The highest burden of child stunting is in Africa and Asia. In Africa, although the prevalence of children under-5 years of age who are stunted is decreasing, the number of children affected by stunted linear growth is actually increasing.

The determinants of undernutrition are complex and interacting. (7,8) Inadequate dietary intakes and inadequate care are important immediate and underlying causes, respectively. The World Health Organization (WHO) indicators of minimum acceptable diet for children 6-23 months of age comprise both dietary diversity and meal frequency. (9) For infants and young children, "care" is comprised of several practices such as food preparation, feeding, psychosocial and cognitive stimulation, hygiene practices, and home health practices. (10) Good childcare takes time, and usually it is women's time, because around the world – despite high rates of women's engagement in income-generating work – the burden of childcare lies primarily on women's shoulders. (11,12)

The most accurate and reliable way to collect diet and time-use data is through direct observation (OBS), which requires a researcher to be at the study participant's home all day, observing and recording everything they eat and do. (12–20) Although OBS is accurate and reliable, it requires a great deal of expertise, skill and motivation on the part of the observer; incurs high time and financial costs; and imposes a heavy burden on participants. (21–23) Having an observer present in the home can also influence participant behaviour, due to various forms of reactivity bias, including the Hawthorne effect, observer-expectancy effect, and observer bias. In lieu of OBS, in low-income countries (LICs), a common method of data collection for diet or time use is the 24-hour recall (24HR). This is a structured interview where the interviewer asks the respondent about all the activities s/he performed and/or foods / beverages consumed the previous day. While less costly and less burdensome to the participant, there are well-documented drawbacks of recall methods in terms of data quality. (13,24–27)

Specifically, as a retrospective method, the quality of data obtained via the 24HR depends on the enumerator's verbal probing skills and persistence, and the respondent’s motivation and capacities to understand the content of the questions asked, accurately recall the information needed, identify the relevant information, and respond in the format requested. Recall methods are also prone to respondent biases such as social desirability bias and approval bias, and researcher biases such as leading questions and wording. Deficits in any of these areas create opportunities for misreporting. Misreporting – be it intentional or unintentional - creates error in the data. For both diet and time use, that error may be underreporting, overreporting or a combination of underreporting and overreporting, depending on the food / beverage or activity.

Automated wearable cameras (AWCs) are inexpensive technologies that prospectively and unobtrusively record activities as they are performed. In high-income countries (HICs), the feasibility and/or validity of an AWC as a research method have been investigated and published in peer-reviewed publications for dietary intake, (28–38) food preparation practices, (39,40) food environment, (32,33,37,41–46) time use, (47–49) physical activity, (45,50–57) and activity space. (55,58–61) Only five of these studies indicate the location type in which the data were collected: two were conducted in urban areas (41,58) and the remaining were conducted in a mix of urban, suburban and rural areas. (56,60,61) These studies have shown that AWCs are overall, from a participant's perspective, feasible and acceptable for use by female and male children, teenagers, adults, and older adults in HICs. Although some participants in six studies reported stress induced by wearing the AWC, (28,34,52,60,62,63) the burden of wearing the device was not a major issue. These studies from HICs do, however, highlight potential feasibility challenges from a researcher's perspective. Loss of data due to device inoperability, (28,29,31,33,34,36,39,41,42,46,52,53,58,59) poor image quality (e.g., dark or blurry images), (33,34,36,41,42,51,52,56,59–61,64–66) and poor camera positioning were commonly reported. (28,34,36,42,43,57,59,64,66) One study reported the cost of purchasing the AWCs, (47) but no studies compared the costs of AWC-based methods to traditional methods overall. Twelve studies reported that the amount of researcher time required to code AWC images was onerous. (31,33,39,41,42,45,47,50,56,61,65)

The three dietary intake validation studies of AWC-based methods conducted in HICs assessed energy intake among adults by coding individual foods / beverages and portion sizes consumed. (34,36,38) Two were pilot studies having ten or fewer participants. (34,38) The results were consistent across the three studies. The AWC-based method resulted in higher reported energy intake compared to 24HR (34,36) and food diary methods, (38) although energy intake still underestimated total energy expenditure. (36,38) All studies found that fewer foods were unreported when using the AWC compared to traditional methods.

The three time use validation studies of AWC-based methods conducted in HICs assessed time allocated to activities among adults. These studies considered the automated wearable camera-based image-assisted recall (AWC-IAR) as the criterion, and the time-use diary (TUD) method as the comparator. Two of the articles analysed data from the same study at different levels of aggregation. (48,49) The studies show that for most activities, differences in mean time allocation estimates were similar between the AWC-based method and the TUD method. Harms et al. (2019) and Gershuny et al. (2020) showed significant differences in mean time allocated to eating. (48,49) Kelly et al. (2015) showed significant time allocation differences for "hobbies/computing" and "media" only. (47) Kelly et al. (2015) also reported the number of activities performed. The AWC-IAR resulted in more reported activities than the TUD method.

Only three AWC feasibility, acceptability and/or validity studies have been undertaken outside of HIC contexts, all of which were conducted in middle-income countries and published after the data for this thesis were collected. AWC have been evaluated for assessing dietary intake in China and Tonga, (67,68) food environment in Tonga, (68) and time use in Ghana. (69) None of these studies expressly assessed the feasibility or acceptability of the AWCs. Veatupu et al. (2019) indicates that some data may have been lost due to device inoperability and that the images were not suitable for assessing the quantity of foods consumed, as initially intended. (68) Foods that are commonly shared, such as snacks, were especially difficult to assess. Only Zhou et al. (2019) reported on the feasibility of AWCs from the researcher perspective, indicating that it took approximately 40 minutes to code the foods and portion sizes in pre-selected images (n=6-8) captured by the AWC. (67) Zhou et al. (2019) also reported on the comparability of an AWC-IAR, as the stated reference method, to 24HR for assessing energy intake among school-age children in China. Compared to the 24HR method, the AWC-IAR method resulted in higher reported energy intakes. Laskaris et al. (2019) reported on the validity of AWCs for assessing the activities performed by e-waste workers in Ghana. (69) The agreement in time-activity patterns between the AWC-based method and the TUD method was poor.

Substantial gaps remain in determining the validity of AWCs as a research method. In particular, no feasibility or validity studies have been conducted in a LIC. Even among the studies conducted in middle- or high-income countries, none have focused on rural contexts or women caregivers of children under two years of age. Key characteristics of rural women residents of LICs, such as literacy, exposure to technology and social norms, are quite different compared to any of the populations targeted in the AWC research conducted thus far. The research environment in rural LICs also poses different challenges including, for example, limited access to electricity for lighting the activity space or charging devices, and higher chance of AWC exposure to dirt or liquids, and lack of enumerators having pertinent skills. Although results of studies investigating AWC-based methods in HICs are promising, research is needed to assess the feasibility and validity of AWC-based methods in rural low-income environments.

CONTEXT

Prior to the COVID-19 pandemic, the prevalence of stunting among children under 5 years of age was decreasing in Uganda, from 45 percent in 2000-01 to 29 percent in 2016. (70) In the Busoga region of Uganda, where this study took place, the prevalence of stunting among children less than five years of age in 2016 was 29 percent. Among all Ugandan women of childbearing age in 2016, 8.7 percent were thin (<18.5 body mass index [BMI]) and nearly a quarter (23.8%) were overweight or obese, whereas in the Busoga region, 6.6 percent of women were thin and 16.9 percent of women were overweight or obese.

In Uganda, the diversity of foods consumed by children aged 6-23 months is low. The 2016 Demographic and Health Survey (DHS) found that less than a third (30.3%) of children 6-23 months of age achieved minimum dietary diversity. (70,71) Dietary diversity was only slightly higher in the Busoga region (31.1%) compared to the national average.

According to the *Uganda Time Use Survey Report 2017-2018*, Ugandan women spent most of their time (approximately 13 hours per day) engaged in so-called unproductive work[1]. (72) The activities and activity groups used by the Uganda Bureau of Statistics (UBOS) are available in Appendix 1. On average, another 6.7 hours per day were spent doing unpaid work[2], including 2.4 hours per day spent doing unpaid domestic services and 0.9 hours per day spent doing unpaid caregiving services. On average, only about 3.5 hours per day were spent doing productive work[3], including employment (2.1 hours) and production of goods for their own final use (1.4 hours). In the Eastern Region (which encompasses the region formerly called Busoga), the amount of time allocated to unproductive work (14.0 hours) and unpaid work (7.5 hours) was higher than the national average.

Aims and Objectives

New research methods that improve data quality while minimizing participant burden are needed for assessing the immediate and underlying drivers of nutrition in LICs. The evidence from HICs indicates that AWCs as a research data collection method are feasible for a variety of populations and research areas and may decrease misreporting of dietary intake and time-use data. There is no evidence, however, regarding the feasibility or validity of AWCs from LICs, rural populations, or

---

[1] Unproductive work comprised socializing and communication, community participation and religious practices; culture, leisure, mass media and sports practices; and sleeping, eating, and drinking, personal hygiene and care, receiving personal health care and travels related to self-care and maintenance.

[2] Unpaid work comprised unpaid domestic services for household and family members; unpaid caregiving services for household and family members; unpaid volunteer, trainee, and other unpaid work; and learning.

[3] Productive work comprised employment and related activities, and production of goods for own final use.

women or children under two years of age. The aim of this thesis was to develop and validate a method using an AWC for assessing maternal and child dietary diversity and women's time use in rural Eastern Region Uganda. Specifically, the objectives of this thesis were to:

1. Assess the feasibility of an automated wearable camera-based image-assisted recall for use in collecting dietary and time-use data in rural Eastern Region Uganda.
2. Evaluate the concurrent criterion validity, for assessing maternal and child dietary diversity, using a semi-quantitative automated wearable camera-based image-assisted recall method and a semi-quantitative multiple-pass 24-hour dietary recall method, compared to weighed food records in rural Eastern Region Uganda.
3. Evaluate the concurrent criterion validity, for assessing women's time use, of an automated wearable camera-based image-assisted recall method and a 24-hour activity recall method, compared to direct observation in rural Eastern Region Uganda.

It was hypothesized that an AWC-IAR method for assessing women's and children's dietary diversity and women's time use was feasible in a rural LIC context and could provide accurate and reliable data compared to OBS, at least not worse than data gathered via 24HR.

Thesis Roadmap

This thesis follows a research paper style, with seven chapters in total. The current chapter introduces and frames the thesis. Chapter 2 is a series of systematic scoping reviews summarizing three areas: (1) studies assessing the feasibility of AWCs for research data collection; (2) validation studies for methods of assessing dietary diversity, and (3) validation studies for methods of assessing women's time use. Chapter 3 provides an overview of the methods used in this thesis. Chapters 4-6 comprise three articles. The first article (Chapter 4), under review with *Nutrition Journal* (as of November 2022), assesses the feasibility of the AWC-IAR method. The second article (Chapter 5), published in the *British Journal of Nutrition*, is a validation of an AWC-IAR method and 24HR method for assessing maternal and child dietary diversity. The third article (Chapter 6), published in *Nutrients*, is a validation of an AWC-IAR method and 24HR method for assessing women's time use. Please note that the articles (Chapters 4-6) are published under my married name, Andrea L. S. Bulungu. The final chapter (Chapter 7) synthesizes key research findings and discusses how the research has advanced global knowledge regarding methods for assessing maternal and child dietary diversity and women's time use.

COLLABORATING INSTITUTIONS AND FUNDING

The collaborating institutions for this research included: London School of Hygiene and Tropical Medicine (LSHTM), the Africa Innovations Institute (AfrII), the University of Greenwich Natural Resources Institute (NRI), and the Sasakawa Africa Association (SAA). Funding for research costs was provided by UK Aid from the UK government and the Bill & Melinda Gates Foundation (BMGF) through the Innovative Methods and Metrics for Agriculture and Nutrition Actions (IMMANA) programme.

ROLE OF THE CANDIDATE

A snapshot of the thesis timeline is presented in Figure 1.1. Please note that I was a part-time student for the duration of the PhD.

Proposal

I made significant contributions to proposals for other studies before finally landing on the study that became the foundation for this thesis. When I joined the IMMANA study team in March of 2017, the overall conceptual design, research questions and objectives for the study were already defined. During this period, the IMMANA team –myself included–developed a proposal in response to the "Drivers of Food Choice" request, which built on the IMMANA study. Aspects of this proposal were incorporated into my upgrading report. I made major contributions to the proposal, which was ultimately funded. My contributions to the IMMANA study began with drafting the research instruments and contributing to the study design. I upgraded in August 2017. In November 2017, I joined the team in Uganda to begin final preparations for data collection.

Preparation

In Uganda, I worked with Elaine, Joweria, Kate, Jan, and Gwen to aggregate and prepare all of the equipment and supplies used in the study[4] (e.g., mobile phones, wearable cameras, and GPS trackers; tablet computers; sim cards, memory cards, and data banks; food weighing scales, anthropometric equipment, etc.) (Figure 1.2) The Information and Communication Technologies (ICTs) (i.e., wearable cameras, GPS trackers, tablet computers, data banks and peripherals), and food weighing scales used in the study were selected and purchased overseas by Jan Priebe (ICTs), Elaine Ferguson (scales) and Kate Wellard, and carried into Uganda by members of the research team. Mobile phones and sim cards (used for the IVR), and memory cards (used by the AWCs) were purchased in Uganda. Anthropometric equipment was borrowed from a Ugandan hospital. During this period, with guidance primarily from Elaine, I continued developing the diet and time-use assessment instruments and protocols for the three methods (OBS and weighed food records, 24-hour [24h] diet and activity recalls, and AWC-IAR), and the anthropometry instrument and protocol.

Towards the end of December, I contributed to the training of 18 data collectors and 3 supervisors for the IMMANA study. With Gwen, I trained data collectors on the mother's questionnaire. I led the training on the semi-quantitative multiple-pass 24h dietary recall method, the ICTs, and the AWC-IAR method. At the end of training, we assessed the data collectors' skills performing all methods. I assessed their competence in handling the ICTs and in facilitating the AWC-IAR. We then had two days of field practice.

Following institutional review board (IRB) approval by both institutions (LSHTM and the Uganda National Council for Science and Technology [UNCST]) in January 2018, we did iterative piloting of the AWCs and AWC-IAR instrument.

Data Collection

Data collection began 11 January 2018. Given the tight timeline between instrument piloting and refinement and data collection, all non-ICT data was collected on paper.

[4] This thesis was part of a broader research study to assess the validity of three innovative methods: an interactive voice response (IVR) system using rudimentary (button) mobile phones, global positioning system (GPS) trackers and the AWC. The focus of the thesis, however, is the AWC-IAR.

In the field, I was responsible for day-to-day study management and data quality control over the 6 weeks of data collection. I assigned teams to their respective communities each week, and each data collector to their respective households each day. Specifically, in the evening, between 9pm and 11pm I received all equipment from the 18 enumerators via their supervisor; checked in and organized all forms; labelled and archived data from each camera and GPS tracker; and charge all devices in preparation for the next day. While the devices and tablets charged, I prepared all data collection tools and device assignments for the following day and packaged all supplies by enumerator identification number (ID). During data collection, therefore, the brunt of my work occurred between 9pm and 3am.

By 5am, the teams were heading to the field for another day of data collection. Gwen distributed the instruments, devices, and requisite supplies to supervisors that I had prepared the night before. During the day, while the teams were in the field, I reviewed the forms we had received the night before and addressed any issues that arose. In the penultimate week of data collection, I missed a few days due to typhoid. Data collection ended on 16 February 2018.

Data Entry and Preparation

Aside from the AWC-captured photos, GPS tracks, and IVR recordings, all study data used in this thesis was collected on paper. In the final week of data collection, I began developing data entry tools using EpiData. By March 2018 I had onboarded two data entry research assistants (RAs) who had been recommended by a colleague. Bernice, Sarah, and I would work together in Kampala for over a year to enter all study data, with the exception of the mother's questionnaire (which was entered by Gwen and Joweria). In general, qualitative data were entered via tools I created in EpiData, whereas more structured data (e.g., the time-use matrices) were more easily entered by the RAs via Excel.

Data entry, preparation and analysis occurred in cycles, by instrument. As soon as the RAs completed data entry for an instrument, I would begin verifying, cleaning, and preparing it for analysis. Once the data was digitized, I worked closely with Elaine and Luigi to refine the data analysis plan and methods. With the exception of the Poverty Probability Index (PPI), which was analysed by Pamela, I did all of the data analysis. Elaine, Luigi, and I met on a fairly regular basis to review and discuss results and determine next steps.

Data Analysis, Manuscript Development, and Dissemination

I started data analysis in May 2018, while the RAs continued with data entry. I began with the feasibility analysis. Preliminary results were accepted as a poster presentation at the Agriculture, Nutrition & Health Academy conference in Ghana in June. Also in June 2018, I was invited to give an "elevator pitch" on the AWC-IAR method at the American Society for Nutrition conference in Boston.

Having realized that the feasibility analysis could not be completed until the dietary diversity and time use data had been cleaned and processed, I began preparing and analysing the dietary diversity data. Validation analyses of the dietary diversity data were carried out from September 2018 through April 2020, when the manuscript for the dietary diversity validation study was submitted for publication. Preparation and analysis of the dietary diversity data was relatively straightforward compared to the time-use data.

Analyses of the time-use data began in November 2019, while the dietary diversity validation manuscript was being finalized. The time-use data was substantially more difficult to work with, in part because the AWC-IAR instrument was structurally different from the instruments used for collecting OBS and 24h activity recall data. To make the AWC-IAR data comparable to the OBS and 24h activity data required a lot of processing. Also, compared to dietary diversity data, there is little precedent for processing time-use data that accounts for concurrent activities. The approach we used for the time use data analysis, which was developed by me with guidance from Elaine and Luigi, is a novel contribution. Validation analyses of the time-use data continued until January 2022, when the manuscript was submitted for publication.

With all dietary diversity and time-use data thoroughly processed, analysis of the feasibility data resumed in July 2021 and continued until May 2022, when the manuscript was submitted for publication.

Publications and Conference Presentations

Publications

- **Bulungu, A. L. S.**, Palla, L., Priebe, J., Forsythe, L., Katic, P., Varley, G., Galinda, B. D., Sarah, N., Nambooze, J., Wellard, K., & Ferguson, E. L. (2022). Validation of an Automated Wearable Camera-Based Image-Assisted Recall Method and the 24-Hour Recall Method for Assessing Women’s Time

Allocation in a Nutritionally Vulnerable Population: The Case of Rural Uganda. Nutrients, 14(9), 1833. https://doi.org/10.3390/nu14091833

- Kimere, N. C., Nambooze, J., Lim, H., **Bulungu, A. L. S.**, Wellard, K., & Ferguson, E. L. (2022). A food-based approach could improve dietary adequacy for 12–23-month-old Eastern Ugandan children. *Maternal & Child Nutrition*. https://doi.org/10.1111/mcn.13311
- **Bulungu, A. L. S.**, Palla, L., Priebe, J., Forsythe, L., Katic, P., Varley, G., Galinda, B. D., Sarah, N., Nambooze, J., Wellard, K., & Ferguson, E. L. (2021). Validation of a life-logging wearable camera method and the 24-h diet recall method for assessing maternal and child dietary diversity. *British Journal of Nutrition*, *125*(11), 1299–1309. https://doi.org/10.1017/S0007114520003530

Conference presentations

- **Bulungu, A. L. S.** Validation of an automated wearable camera-based image-assisted recall method and the 24-hour recall method for assessing women's time allocation in Uganda. 44th International Association for Time Use Research Conference. Montréal, Quebec, Canada. 17-19 August 2022. [Oral]
- **Bulungu, A. L. S.** Validation of an automated wearable camera-based image-assisted recall method and the 24-hour recall method for assessing women's time allocation in rural Uganda. 117th American Sociological Association Annual Meeting. Los Angeles, California. 5-9 August 2022. [Oral]
- **Bulungu, A. L. S.**, Ferguson, E. L. Validation of a life-logging wearable camera method and the 24-hour diet recall methods for assessing maternal and child dietary diversity. 4th Annual Agriculture, Nutrition & Health (ANH) Academy Week. Hyderabad, India. 2019. [Poster and Elevator Pitch]
- **Bulungu, A. L. S.** The Feasibility and Acceptability of Information and Communication Technologies (ICTs) for Collection of Diet and Time Use Data. 3rd Annual Agriculture, Nutrition & Health (ANH) Academy Week. Accra, Ghana. 2018. [Poster]
- **Bulungu, A. L. S.** ICTs for Nutrition and Time Use (INATU). Sight and Life Elevator Pitch Contest at the American Society for Nutrition (ASN) Annual Meeting. Boston, MA. 2018. [Elevator Pitch]

Invited presentations

- Bulungu, A. L. S. Validation of an automated wearable camera-based image-assisted recall method and the 24-hour recall method for assessing dietary diversity and time allocation in rural Uganda. Scuola Specializzazione in Statistica Sanitaria e Biometria, Sapienza Università di Roma, 16 March 2022. [Oral]
- **Bulungu, A. L. S.** Assessing maternal and child dietary diversity and women's time use using automated wearable cameras in Uganda. Drivers of Food Choice and Innovative Methods and Metrics for Agriculture and Nutrition Actions Stakeholder Dissemination Workshop, Kampala, Uganda, 2 December 2021. [Oral]
- **Bulungu, A. L. S.** Assessing maternal and child dietary diversity and women's time use using automated wearable cameras in Uganda. Program in International Nutrition, Cornell University, Ithaca, New York, USA, 7 October 2021. [Oral]
- **Bulungu, A. L. S.** Validation of a life-logging wearable camera method and the 24-hour recall method for assessing dietary diversity and time allocation. Scuola Specializzazione in Statistica Sanitaria e Biometria, Sapienza Università di Roma, 15 June 2021. [Oral]

Awards

- Sight and Life Elevator Pitch Contest 3rd Place Award (2018)

Figure 1.1. Thesis timeline

Please note that I was a part-time student for the duration of my PhD.

	2014	2015	2016											
			J	F	M	A	M	J	J	A	S	C	N	D
Key PhD Milestone			X			X								
Application	X	X												
Project proposal	X	X		X	X	X	X	X	X					
Project preparation					X	X	X	X	X	X	X	X	X	X
Data collection										X	X	X	X	X
Data entry														
Data preparation														
Data analysis														
Dissemination														

	2017											
	J	F	M	A	M	J	J	A	S	O	N	D
Key PhD Milestone								X				
Application												
Project proposal							X		X		X	
Project preparation	X	X	X	X	X	X	X	X	X	X	X	X
Data collection										X	X	X
Data entry												
Data preparation												
Data analysis												
Dissemination												

	2018											
	J	F	M	A	M	J	J	A	S	O	N	D
Key PhD Milestone												
Application												
Project proposal												
Project preparation	X											
Data collection	X	X										
Data entry		X	X	X	X		X		X	X	X	X
Data preparation			X	X	X	X	X	X	X	X	X	X
Data analysis					X	X	X	X	X			
Dissemination				X	X	X						

	2019											
	J	F	M	A	M	J	J	A	S	O	N	D
Key PhD Milestone												
Application												
Project proposal												
Project preparation												
Data collection												
Data entry	X	X	X									
Data preparation	X	X	X	X	X	X	X	X	X		X	
Data analysis	X	X	X	X	X	X	X	X	X	X	X	
Dissemination			X	X	X	X	X	X	X	X	X	X

	2020											
	J	F	M	A	M	J	J	A	S	O	N	D
Key PhD Milestone												
Application												
Project proposal												
Project preparation												
Data collection												
Data entry		X										
Data preparation	X	X	X					X			X	X
Data analysis	X	X	X	X	X	X	X	X	X	X	X	X
Dissemination		X	X	X	X	X	X	X	X			

	2021											
	J	F	M	A	M	J	J	A	S	O	N	D
Key PhD Milestone												
Application												
Project proposal												
Project preparation												
Data collection												
Data entry												
Data preparation	X									X		
Data analysis	X	X	X			X	X	X	X		X	X
Dissemination	X	X	X			X	X	X	X	X	X	X

	2022											
	J	F	M	A	M	J	J	A	S	O	N	D
Key PhD Milestone										X		
Application												
Project proposal												
Project preparation												
Data collection												
Data entry												
Data preparation												
Data analysis	X	X	X	X								
Dissemination	X				X	X	X	X				

Figure 1.2 ICTs used in the study

Chapter 2: AWC-based IAR and 24HR methods for assessing maternal and child dietary diversity and women's time use - Three systematic scoping reviews

INTRODUCTION

Globally, despite improvement, both the burden and cost of undernutrition remain high. Nearly one in five children globally are stunted. (6) Stunting carries increased risk for illness, early death, and poor development, which is ultimately measurable in terms of lost earnings and lost GDP at the country level. (1,4,5) Although globally the prevalence of stunting is decreasing, the highest burden of child stunting has been and remains in Africa and Asia. Africa is the only region where, although the prevalence of stunting among children under 5-years of age is decreasing, the number of children affected by stunting is increasing.

The causes of malnutrition are numerous and interacting. (7,8) The immediate and underlying drivers include poor dietary intake and inadequate care. Increasing the diversity of nutritious food groups in the diets of children aged 6 to 23 months, increases the likelihood of dietary adequacy, (73–76) and dietary diversity has been positively associated with linear growth. (77,78) Child care is comprised of several practices such as food preparation, feeding, psychosocial and cognitive stimulation, and hygiene and other home health practices. (10) The provision of quality childcare relies on caregivers' own health and nutritional status, and time availability. Around the world, women remain primarily responsible for food preparation and childcare. (79) Women's time availability is therefore likely to be an important factor related to the uptake of good diet and care practices for themselves and their children.

Popkin (1980) first demonstrated that a mother's time spent on childcare was positively associated with child nutrition status. (80) Time for care was subsequently included as a key factor driving maternal and child nutrition and overall well-being in the Engle et al. (1987) model of care. (10) Yet, decades of empirical studies have shown that the relationship between women's time allocation and maternal and/or child nutrition to be complex. Gryoboski (1996) found that time allocated to childcare by aunts, sisters, and grandmothers was associated with increased caloric intake for children, (81) whereas the association between the mother's time allocated to childcare and caloric intake was negative, and Komatsu et al. (2018) found that the association between women's time allocated to "reproductive work" and child nutrition depended on the economic status of the household. (82) Other studies have taken the opposite approach, exploring instead the influence of women’s time allocated to productive work on their and/or their child’s nutritional status. The results of these studies have also been conflicting. Some have found that a mother’s time allocated to agriculture work was negatively associated with child nutritional status, (83,84) whereas others have found there to be no relationship between a mother’s time allocated to work and child nutritional status. (85,86)

Previous studies indicate that the amount of time women allocate to both productive and reproductive work is severely undercounted, (16,87–92) which limits our ability to accurately assess how women spend their time and its influence on women’s and children’s nutrition, health, and well-being. (83,93–100) Accurate and timely maternal and child dietary diversity and women's time-use data are needed to understand relationships between women’s time use and child/maternal dietary quality, and to identify populations at risk of undernutrition due to the consumption of poor-quality diets. Such data are important for designing appropriate interventions and evaluating change in dietary quality or time-use patterns.

The most accurate and reliable method of collecting diet and time-use data is OBS. (12–20) Although OBS is accurate and reliable, it requires a great deal of expertise, incurs high time and financial costs, may influence participant behaviour due to the presence of the observer, and imposes a heavy burden on participants and the research team. (21–23) In lieu of OBS, in LICs, a common method of data collection for dietary intake or time-use data is via 24HR methods. The 24HR method is recommended by the Food and Agriculture Organization of the United Nations (FAO), WHO and United Nations Children's Fund (UNICEF) for assessing the dietary diversity score (DDS) of women and children, respectively. (9,101) The 24HR method is also used in the time-use modules of the Women's Empowerment in Agriculture Index (WEAI) and the Living Standards Measurement Study (LSMS). (102,103)

While less costly and less burdensome to the participant, there are well-documented drawbacks of recall methods in terms of data quality. (13,24–27) Specifically, as a retrospective method, the quality of data obtained via 24HR depends on the enumerator's verbal probing skills and persistence, and the respondent’s motivation and capacities to understand the content of the questions asked, accurately recall the information needed, identify the relevant information, and respond in the format requested. Recall methods are also prone to respondent biases such as social desirability bias and approval bias, and researcher biases such as leading questions and wording. Deficits in any of these areas create opportunities for misreporting. Misreporting, whether it is intentional or unintentional, creates error in the data.

AWCs are inexpensive technologies that have been used in HICs to prospectively collect health-related data, including data on dietary intakes and time use. (28,29,31–38,47–49) The images captured by the AWC must be later coded for analysis. To date, images captured by AWCs have been coded in three ways: (1) via an image assisted recall (IAR) wherein the AWC images are used to "trigger" participants' memories, (2) by a researcher without the input of the study participant (referred to in this thesis as an enumerator image interpretation), and (3) via artificial intelligence (AI).

There are numerous studies that have been undertaken using 24HR and AWC-based methods for assessing dietary intake and time use, broadly speaking. The purpose of these systematic scoping reviews was to summarize the evidence regarding the feasibility of AWC-based research methods generally, and the validity of the 24HR and AWC-IAR, specifically, for assessing maternal and child dietary diversity and women's time use, with particular interest on findings pertinent to rural LIC contexts. (104) That is, regarding the feasibility of AWC-based research methods (regardless of the method used for coding the data), this review aimed to determine: (1) where AWC-based methods have been evaluated for feasibility; (2) for which populations, and for which health-related outcomes they have been assessed; and (3) the major feasibility issues identified. Regarding the validity of the 24HR and AWC-IAR (i.e., specifically using an image-assisted recall to code images captured by an AWC) for assessing dietary diversity and time use, this review aimed to determine: (1) the countries and contexts where the 24HR and/or AWC-IAR methods been used to assess dietary diversity and/or time use; (2) for which populations they have been validated; and (3) the validity of these two methods, for assessing dietary diversity and/or time use, in these different contexts and populations.

SEARCH STRATEGY

CINAHL Plus, EMBASE, Global Health, MEDLINE, PsycInfo, PubMed, Scopus, and Web of Science databases were searched electronically without date restrictions up to January 2021. The systematic search was conducted between September and October 2021 using the search terms presented in Figure 2.1, Figure 2.3, and Figure 2.5. These terms reflected the methods, outcomes, and indicators of interest to identify relevant studies. To maximize the capture of relevant studies, no exclusion keywords were used. Manual searches of personal files and reference lists of identified studies were also conducted to supplement the search results.

Inclusion and Exclusion Criteria

To be included in the review, studies were required to meet the following criteria: (1) published in the English language; (2) published in a peer-reviewed journal (3) open access or otherwise available to ALSB through the LSHTM library's special access rights or shared by the author(s); and (4) human subjects. Published abstracts (e.g., from conferences) without full text were excluded. Literature reviews were also excluded to avoid duplication and ensure that reported results fit the criteria of this study, although their references were reviewed for any pertinent studies not otherwise identified.

Identification of relevant studies

Search results were aggregated and analysed in Excel, for flexibility and ease-of-use. Duplicates from across the eight databases were removed. The remaining studies were then screened iteratively to determine eligibility based on the above inclusion and exclusion criteria. The article titles and abstracts were first manually screened and those which did not meet the inclusion criteria were excluded. The remaining full-text articles were then reviewed for eligibility. Due to resource constraints, article screening was performed by one person (ALSB). Unclear cases were retained until they could be conclusively excluded or included.

Extraction and Synthesis of Evidence

The following parameters were extracted from each eligible article: location; outcome of interest and objective; target population; sample size and sampling method; and key findings.

Scoping Review 1: Feasibility of AWC-based research methods

Search Strategy

Figure 2.1 presents the search terms used in the search for AWC-based methods feasibility studies. The first group included words related to the automated wearable camera, including brand names of specific wearable cameras commonly used in AWC research. The second group of search terms aimed to capture studies that had explored factors related to feasibility. The third group aimed to capture studies using the AWC as a method of research data collection, rather than as an intervention (e.g., a tool to aid memory for individuals suffering from cognitive disorders).

Figure 2.1. Search terms for feasibility of AWC-based research methods literature search

1	wearable camera or wearable cameras or automated imaging or automated image-capture or life-logging or lifelogging or SenseCam or Autographer or eButton or Vicon Revue or micro-camera or narrative clip
2	feasibility or acceptability or usability
3	assess or assessment or assessing or research or researching or data collection
4	1 and 2 and 3

Inclusion and Exclusion Criteria

In addition to the inclusion and exclusion criteria applied to all three scoping reviews described above, the feasibility search included studies using both qualitative and quantitative methods. Studies with the following characteristics were excluded: (1) the AWC was assessed for a purpose other than research data collection (e.g., as a clinical intervention such as fall prevention or memory rehabilitation); (2) the outcome of interest was unrelated to feasibility; (3) the AWC was not actually worn by study participants (e.g., focus group discussions about the perception of AWC); (4) the camera was not automated (e.g., food photography); or (5) the camera did not take still photographs (e.g., wearable video cameras).

Extraction and Synthesis of Evidence

In addition to the extraction and synthesis of evidence steps described above, information was also extracted regarding the brand of AWC device used and its image capture frequency for the articles related to AWC feasibility. Results were tabulated and summarized by outcome of interest, research setting (i.e., level of economic development per the World Bank income classifications[5]), and issue (participant-, community-, or researcher-related).

Results

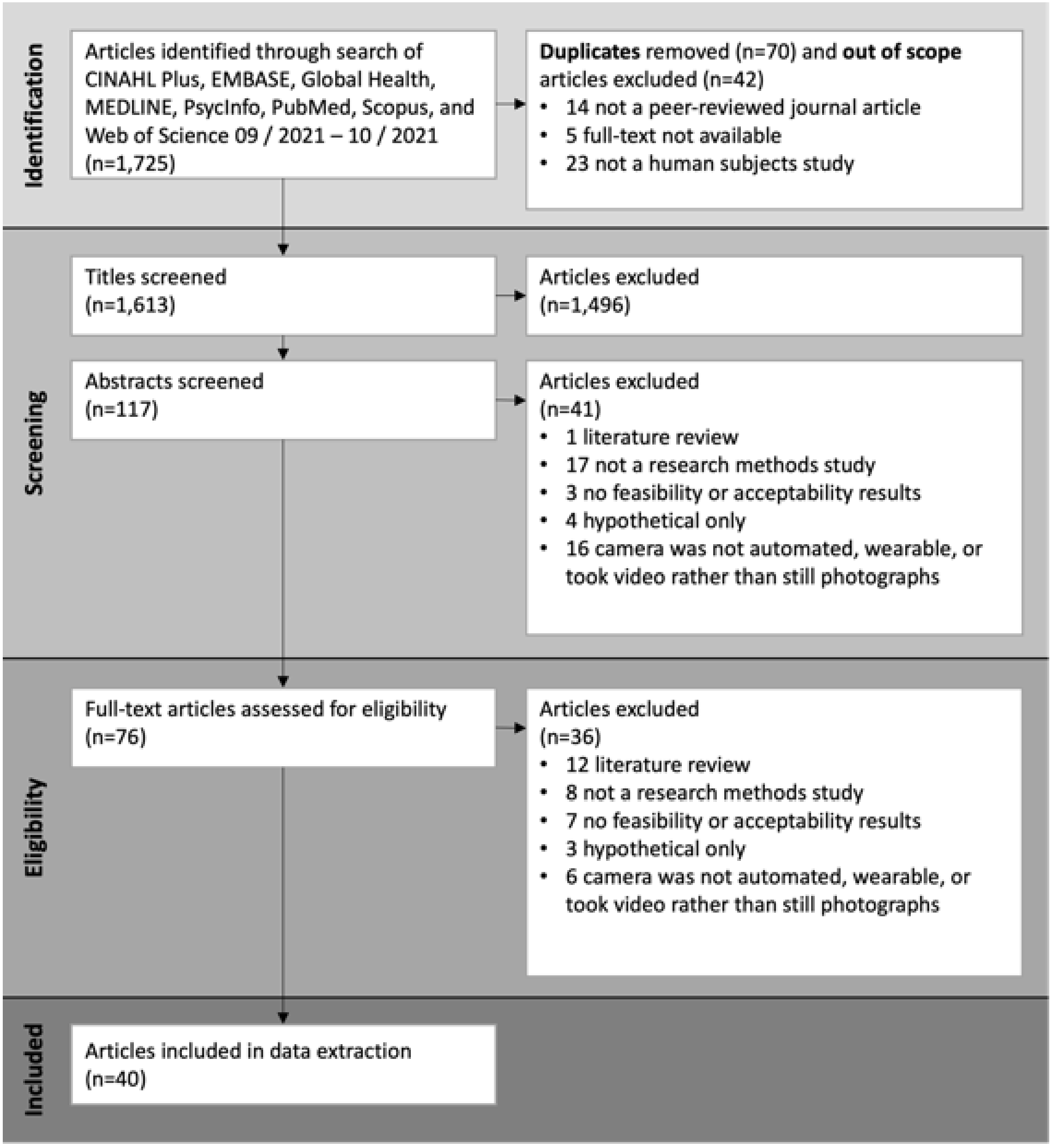
For the feasibility studies, the initial search yielded 1,725 studies, which were reduced to 40 after removing duplicates, screening the titles, and screening the abstracts. (Figure 2.2) Of the feasibility studies included in the review, 60% were identified by Web of Science, 30% by PubMed, and 20% by Medline. Ten articles were found by more than one search engine. In addition, there were five articles included in the review that came from the author's (ALSB) personal files which were neither found by any of the search engines nor included in the reference lists of other articles. For three of these studies, despite using wearable cameras as a research method, did not include any variation of "wearable camera" in their keywords. It is

---

[5] The World Bank categories countries into four income categories: low income , lower-middle-income, upper-middle income, and high income. (156)

unclear why the other two, which did include "wearable camera" as a keyword, were not picked up by any of the search engines. There were no feasibility studies conducted in LICs.

Figure 2.2. Flow diagram of study selection for the feasibility of AWC-based research methods literature search



Description of Studies

In total, 40 studies were found that examined the feasibility of using an AWC to capture data. (Table 2.1) The studies focused on a variety of research areas, including dietary assessment (11 studies), the food environment (10 studies), food preparation practices (2 studies), time use (3 studies), physical activity (7 studies), and activity space (5 studies). Five studies covered multiple research areas (e.g., dietary intake and the food environment). Seven studies addressed other research areas, such as sun protective behaviours, exposure to blue space, and screen use.

Included studies were conducted in Australia (n=1), China (n=1), Ghana (n=1), Ireland (n=2), New Zealand (n=15), Tonga (n=1), UK (n=12), and USA (n=9). A few studies were conducted in multiple countries. Among the included studies, 18 were conducted with adults, (28–30,34,36,40,42,45,47,48,50,51,53,55,56,59,61,62,66) 15 with children under the age of 13 years, (31,32,37,39,41,43,44,46,52,58,64,65,67,68,105) eight with adolescents, (30,32,33,37,39,43,57,60) and three with older adults. (54,62,63) Some studies included participants from multiple age groups. One study was conducted with smokers, (66) three were conducted with university employees, (51,56,61) two were conducted with athletes, (30,56) two were conducted with parents, (40,45) one was conducted with e-waste workers, (69) and one was conducted with people having chronic pain. (63)

The feasibility studies included in this review assessed a variety of AWC devices. Seventeen used SenseCam, (30,32–34,36,41,42,45,51,54–56,59–63) eleven studies used Autographer, (37,46–48,50,52,58,64,65,68,105) three used eButton, (31,39,40) and one study each used Brinno TLC, (57) GoPro, (69) and Narrative Clip. (67) In addition, four studies used a bespoke AWC (28,29,53,66) and for two studies the AWC was unspecified. (43,44)

The included studies covered a range of issues related to feasibility, which can be categorized by audience, i.e., from a participant perspective (33 studies), from a community perspective (10 studies), and from a researcher perspective (38 studies). Most studies reported on feasibility issues from multiple perspectives.

Feasibility from a participant perspective

All the studies reporting on AWC usability (n=21) were conducted in HICs. Among them, ten studies indicated that the AWC may be cumbersome to wear, especially during physical activity. (28,31,32,45,47,52,54,57,62,63) Studies in which participants were responsible for operating the device (e.g., turning the AWC on and off at the start and end of the data collection day) commonly reported that participants forgot to wear or charge the AWC, (45,55,57,59,60,67,105) or had difficulty pressing the buttons on the device. (30,51,54,63) These issues occurred across all age groups. In seven studies (all conducted in HICs or upper-middle-income countries [UMICs]), adults (28,30,34,47) and adolescents (30,32,33) and children (32,67) reported that viewing the images from the AWC helped them to recall pertinent details during the IAR.

Eighteen studies reported on the emotional burden of wearing the device. The single study conducted in a LMIC (69) indicated that the AWC reduced participant burden compared to a traditional TUD. In HICs, six studies indicated that a few participants felt emotional discomfort as a result of wearing the device, especially in public. (28,34,52,60,62,63). Nine studies (conducted in either HICs or UMICs) reported the study participation rates, which ranged from 16 to 100%. (28,31–33,45,52,59) Four studies, with refusal rates ranging from 11 to 84%, explicitly attributed refusal to participate in the study to the AWCs. (33,45,59,63) Pertinent reasons given by individuals who refused to participate included concerns about wearing the device in the workplace, invasion of privacy, physical or emotional discomfort, fear of breaking the device, and fear of being unable to operate the device.

Study withdrawal (31,32,47,59) and non-compliance (29,31,34,36,37,45,47,48,50–54,56,57,59–63,67,68,105) were also reported. Three studies (all conducted in HICs with either children or the elderly) attributed withdrawal to the AWC. (31,32,63) Explanations for dropping out of the study were fear of operating the device, privacy concerns, and "negative attention". Participants in nine studies reported removing the camera while at their child's school; preparing their child for bed; reading confidential information; resting or at leisure; performing what they perceived to be "boring" or repetitive activities; playing in sports; at the doctor's office; or when others were made uncomfortable by the device, or due to fear of operating the device incorrectly. (32,41,45,47,48,52,57,62,63) Study participants also reported removing the device while practicing personal care activities such as toileting, however this would not ordinarily be categorized as non-compliance, per se.

In accordance with the ethical framework outlined by Kelly et al. for AWC in health behaviour research, (106) most studies provided participants the opportunity to review the images captured by the AWC privately, before being viewed by the researchers, and to delete any images they desired. Among the seven studies reporting on the acceptability of the content captured by the AWC (all in HICs), six reported participants opting to delete images. (46,52,56,57,62,63)

Only two studies (both in HICs) reported, specifically, on the acceptability of the AWC-IAR method, (28,47) both indicating participant satisfaction.

Feasibility from a community perspective

Ten studies (all in HICs) indirectly reported on issues related to community acceptability of AWCs worn by study participants. (32,41,47,48,52,54,55,57,62,63) Study participants reported removing or covering the AWC at school, (32,41,47,48) work, (47) home, (52,57) the doctor's office, (48) and in public. (62) Three studies reported participants being approached about the AWC by members of the public without any request to remove the device. (54,55,63)

Feasibility from a researcher perspective

One of the biggest challenges to the use of AWCs as a research method is lost data due to device inoperability (e.g., insufficient battery life or another malfunction). Lost data due to AWC inoperability was reported in 17 studies spanning various country contexts (HIC, UMIC and LMIC) and populations (adults, adolescents, and children). (28–31,33,34,36,39,41,42,46,52,53,57–59,68,69) Data losses, as a proportion of intended image capture, is commonly 15-20%, but can be as high as a third.

Regardless of public response, five studies (all in HICs) reported hyper-awareness of the AWC among participants (33,34,41,45,63) and in six studies (also all in HICs) participants reported having modified their behaviour in reaction to being recorded. (28,29,34,41,45,52)

Most studies reported that the images generated by the AWC were of sufficient quality to enable analysis. However, several image quality issues were common across all contexts (HIC, UMIC and LMIC), including sub-optimal camera angle / positioning; inadequate image capture frequency or otherwise off-camera key events; (28,33,34,36,42,43,57,59,68) dark images caused by low or artificial lighting or an obscured lens; (30–32,34,36,41,42,51,52,54–56,59–61,65,66) and blurry or scrambled images (42,58,65)–all of which may contribute to data loss. The proportion of AWC images deemed "uncodable" ranged from 4-35%.

A few studies also mentioned that AWC images were not suitable for very detailed analyses such as determining the quantity of food consumed, (68) specific items of clothing worn by children far away from the camera, (64) or detecting low intensity activities (e.g., fidgeting or activities performed while sitting down). (50)

The use of AWC aims to maximize reporting accuracy while minimizing participant burden. Part of this burden is shifted to the research team. Many studies (all in HICs) highlight the heavy time burden required to process (41,54) or manually code the AWC images for analysis, (31,33,39,41,42,45,47,50,56,61,65) and susceptibility to human error. (45,56) Not all studies quantified the amount time entailed but estimates of the time required to code

AWC images ranged from approximately 1 to 2 hours per participant day. (39–42,44,46,47,51,56,57,59,60,64,65,67)

Little information on other costs of AWC-based research is available. Only Kelly et al. (2015) reported on the cost of the device (Autographer, £300 each), indicating it was "resource intensive". (47)

Discussion

The feasibility of AWC-based research methods have been evaluated in HICs (Australia, Ireland, New Zealand, the UK, and the USA), UMICs (China and Tonga), and one LMIC (Ghana). The majority have been conducted in just three countries: the UK, New Zealand, and the USA. AWCs were used for collecting data related to dietary intake, the food environment, food preparation, time use, physical activity, and activity space, among others. Most of the feasibility studies have targeted children or adolescents, although studies have also been conducted with adults and the elderly.

For participants, wearing the AWC can be cumbersome, especially during physical activity, and emotionally burdensome. In studies where participants were responsible for operating the device, they sometimes forgot to wear or charge it, and some participants had difficulty pressing the AWC's small buttons. Although several studies reported rates of refusal to participate, withdrawal and non-compliance, only a few concluded that the AWC was a contributing factor. Participants reported removing or covering the device during activities such as childcare, work, and leisure, and some participants reported modifying their behaviours because of wearing the AWC. When given the opportunity, in the studies in which it is reported, some participants opted to delete at least a few AWC-captured images. For researchers, the biggest AWC-related feasibility issues were data losses due to device malfunction, poor image quality, and the heavy resource burden required to process and analyse the data.

There was considerable variability in the breadth of feasibility issues reported. Few studies (n=8) reported on feasibility from the community perspective. Most of the included studies concluded that AWCs were feasible for their respective populations and outcomes of interest, especially when used as an objective measure to complement traditional data collection methods, rather than to replace them. Several studies, however, highlighted major feasibility challenges, such as activities of interest to the research that were missed due to inadequate image capture frequency; device malfunction, which can result in high data losses; ethical concerns regarding privacy; and distress among community members of study participants;

and the time required for coding AWC-captured images. Davies et al. (2020), in particular, concluded that the benefits of AWCs notwithstanding, they are infeasible for large-scale studies. (50)

The period for which participants need to wear the AWC would likely influence feasibility, so it is notable that the evaluation period ranged from less than one day to seven days for the included studies. It may be useful for a future study to assess associations between AWC wear-time and measures of feasibility. Gemming et al. (2015) and Wilson (2016) indicate decreased concerns regarding AWC acceptability over time, whereas feasibility challenges such as forgetting to charge or power the device, or device malfunction may increase with time. (36,63)

Caution is warranted, however, in considering the implications of these studies for a rural, LIC context. Sample sizes varied widely (from 6 to 168), with 29 of the forty included studies having fewer than fifty participants, primarily recruited through convenience or purposive sampling. Most of the studies were conducted with children and/or adolescents, who are more likely than adults to be familiar with and comfortable using life-logging technologies. Yet, even among these high-income youth populations, refusal rates were sometimes 50% or higher. Feasibility issues may be more prevalent and/or severe among a less tech-savvy population.

SCOPING REVIEW 2: VALIDITY OF THE 24HR AND AWC-IAR FOR ASSESSING DIETARY DIVERSITY

Search Strategy

Figure 2.3 presents the search terms used in the search for dietary diversity validation studies. The first group included words related to the 24HR method and the second group to words related to the AWC-IAR method. The third group of search terms aimed to capture studies that had explored factors related to validity. The fourth group included words related to dietary diversity as an outcome of interest.

Figure 2.3. Search terms for the dietary diversity validation literature search

1	24 hour recall or 24-hour recall or 24-hr recall or 24HR or 24-h recall or 24 h recall or 24 hr recall or recall
2	wearable camera or image-assisted recall or image assisted recall or IAR or image-based
3	validation or validity or reliability or accuracy
4	diet quality or dietary diversity or diet diversity or dietary diversity score or diet diversity score or DDS or dietary or food intake or food group intake
5	(1 or 2) and 3 and 4

Inclusion and Exclusion Criteria

For the dietary diversity criterion validity search, in addition to the inclusion and exclusion criteria applied to all three scoping reviews described above, validity studies were excluded if they did not assess concurrent criterion validity of either the 24HR or AWC-IAR methods. Studies were also excluded if the outcome of interest was not individual dietary diversity; the reference method was not either a food record or OBS (i.e., "gold standards" for dietary diversity); or the methods were administered exclusively to populations likely to have increased awareness of dietary consumption (e.g., elite athletes), people with disordered eating (e.g., bulimia nervosa, anorexia), people with cognitive disabilities (e.g., autism) or serious diseases (e.g., kidney disease, Huntington's disease, HIV/AIDS, cancer), or children below 13 years of age.

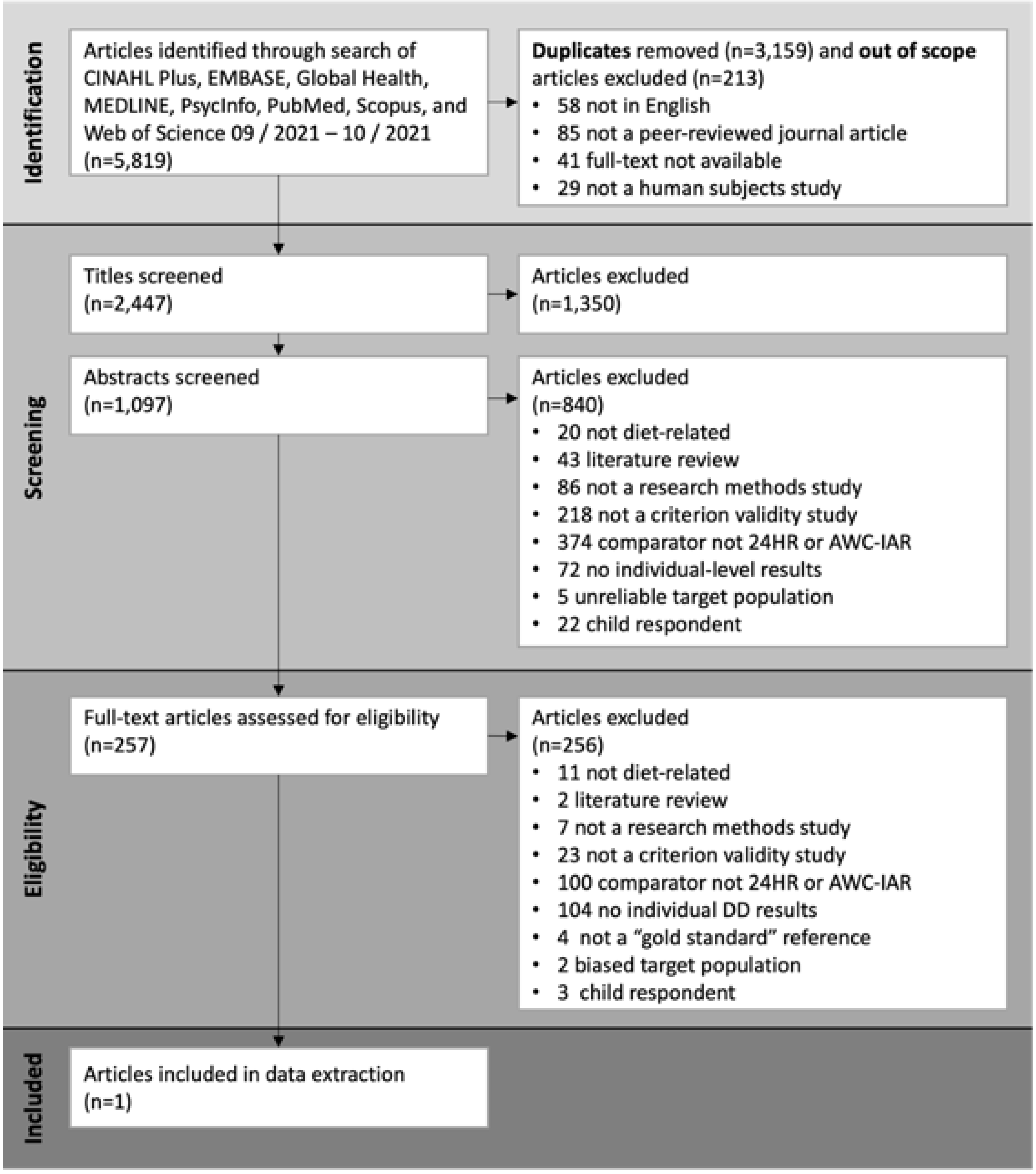
Extraction and Synthesis of Evidence

In addition to the extraction and synthesis of evidence steps described above, information was also extracted regarding the reference and comparator method(s), and number of days of validation for the validation articles. Results were tabulated and summarized by outcome of interest, research setting, and comparator method (24HR or AWC-IAR).

Results

The initial search yielded 5,819 studies, which were reduced to a single dietary diversity study. (Figure 2.4) This study was identified in six of the eight search engines (CINAHL Plus, Embase, Global Health, Medline, PubMed, and Web of Science). The lone dietary diversity criterion validity study pooled data from a LIC (Ethiopia) in addition to two lower-middle-income countries (LMICs) (Cambodia and Zambia).

Figure 2.4. Flow diagram of study selection for the dietary diversity validation literature search



Description of Studies

One criterion validity study for assessing dietary diversity was found. (107) (Table 2.2) This study used the 24HR method to assess the DDS and minimum dietary diversity for women of reproductive age (MDD-W). There were no studies found that assessed dietary diversity using an AWC-IAR method. The one validation study found pooled findings from one LIC (Ethiopia, n=431) and two LMIC (Cambodia, n=430 and Zambia, n=476), comparing the validity of list-based 24HR and open 24HR using weighed food records (WFR) as the reference method. Compared to the WFR, both 24HR methods (list-based and open recall) significantly overestimated DDS (both $p < 0.001$) and MDD-W (30% for WFR vs 46% via the list-based 24HR method and 40% via the open 24HR method, both $p < 0.001$). The reliability of both list-based and open 24HR compared to WFR, measured via Cohen's kappa, were moderate (0.50 and 0.55 for the list-based and open recall, respectively). The food groups most likely to be misreported using 24HR were beans and peas, dark green leafy vegetables, vitamin A-rich fruit and vegetables, and other fruits.

Discussion

Regarding the validity of AWC-IAR or 24HR methods for assessing dietary diversity, the most striking result of this review is the overall lack of evidence. There is only one concurrent criterion validity study of the 24HR method, and none of the AWC-IAR method, for assessing dietary diversity. The 24HR study pooled data for women of reproductive age from Cambodia and Zambia, both LMICs, and Ethiopia, a LIC, which resulted in a large total sample size (n=1,337). The rigor with which this study was undertaken lends credibility to its finding that, compared to the WFR method, the 24HR method significantly overestimated DDS and MDD-W. The implication is that, in similar contexts, when assessing dietary diversity using a 24HR method, women's diet quality would appear better than it actually is.

SCOPING REVIEW 3: VALIDITY OF THE 24HR AND AWC-IAR FOR ASSESSING TIME-USE

Search Strategy

Figure 2.5 presents the search terms used in the search for dietary diversity validation studies. The first group included words related to the 24HR method and the second group to words related to the AWC-IAR method. The third group of search terms aimed to capture studies that had explored factors related to validity. The fourth group included words related to time-use as an outcome of interest.

Figure 2.5. Search terms for the time-use validation literature search

1	24 hour recall or 24-hour recall or 24-hr recall or 24HR or 24-h recall or 24 h recall or 24 hr recall or recall
2	wearable camera or image-assisted recall or image assisted recall or IAR or image-based
3	validation or validity or reliability or accuracy
4	time-allocation or time-use or "time use"
5	(1 or 2) and 3 and 4

Inclusion and Exclusion Criteria

For the time-use criterion validity search, studies were excluded if they did not assess concurrent criterion validity of either the 24HR or AWC-IAR methods.

Studies were also excluded if the outcome of interest was not individual time use (in minutes or hours); the reference was not either TUDs or OBS (i.e., "gold standards" for time use); or the methods were administered exclusively to populations having cognitive disabilities (e.g., autism), or to children below 13 years of age.

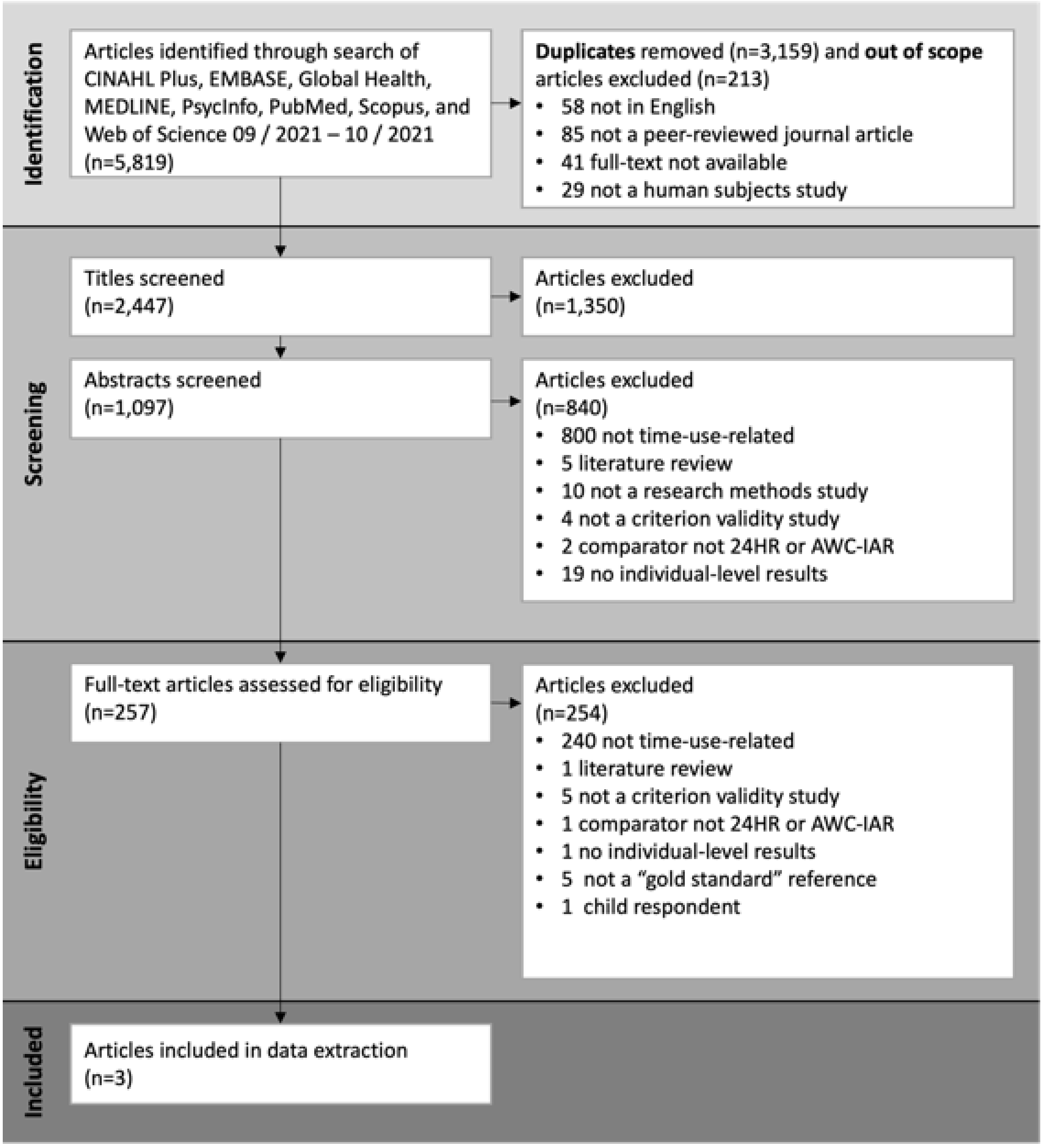
Extraction and Synthesis of Evidence

In addition to the extraction and synthesis of evidence steps described above, information was also extracted regarding the reference and comparator method(s), and number of days of validation for the validation articles. Results were tabulated and summarized by outcome of interest, research setting, and comparator method (24HR or AWC-IAR).

Results

For the validation studies, the initial search yielded 5,819 studies, which were reduced to 3 time use studies. (Figure 2.6) Two of the included studies were identified in Web of Science, one of which was also identified in CINAHL Plus, Embase, Global Health, Medline, and PubMed. One article (AWC-related) included in the review came from the author's (ALSB) personal files; it was neither found by any of the search engines nor included in the reference lists of other articles. It is unclear why this study, which did include "wearable camera" as a keyword, was not picked up by any of the search engines. No time use criterion validity studies were conducted in a LIC, although one was conducted in a LMIC (Ghana).

Figure 2.6. Flow diagram of study selection for the time-use validation literature search



Description of Studies

Three criterion validity studies for assessing time use were found. (48,49,69) (Table 2.3) Two studies, from the CAPTURE24 project conducted in the UK, validated the AWC-IAR for assessing time use compared to the TUD method. (48,49) Notably, these studies considered the AWC-IAR method to be the reference method. They were retained because the results are pertinent regardless of whether AWC-IAR or TUD are considered the reference. The period of validation was the waking hours of one day. The third time use validation study was conducted in a LMIC (Ghana) using TUD as the reference method. (69) The period of validation for this study was a partial work shift, i.e., less than one day. No criterion validity studies for assessing time use using the 24HR method were found.

The two CAPTURE24 studies were administered to adult men and women, whereas 16-50 year old male workers at an e-waste site participated in the Ghanaian study. (69) The CAPTURE24 studies analysed participants' time allocated to activities of the Harmonised European Time Use Surveys (HETUS) framework in 10-minute timeslots using an Autographer camera that captured images at 20-30 s intervals. (108) Laskaris et al. (2019) analysed participants' time allocated to work-related tasks for a portion of the workday in 30-minute timeslots using a GoPro camera that captured an image once per minute. (69) (See Table 2.3 for the list of activities tracked by each study.) The two CAPTURE24 studies built on the Kelly et al. (2015) pilot study, which was included in the feasibility review. (47)

Harms et al. (2019) (n=131) compared time allocations at the HETUS "single-digit" level, comprising ten activity groups (sleep, personal care; eating and drinking; paid work and related; unpaid work, childcare; voluntary/civic activity; social activity, relaxation; physical activity; games, hobbies; media: tv, radio, read, it; travel: all types). (49) Time allocation differences between AWC-IAR and TUD were less than 10% for all but two activity groups. Physical activity (mean time: 25 m vs 21 m for TUD and AWC-IAR, respectively) and eating (mean time: 72 m vs 54 m for TUD and AWC-IAR, respectively) were underestimated by the AWC-IAR method compared to the TUD method. There were statistically significant differences, using pairwise t-tests, between AWC-IAR- and TUD-derived time allocations for sleep, personal care (p=.004); eating and drinking (p <.001); paid work and related (p=.025); and travel (p=.002).

Drawing on the same dataset, Gershuny et al. (2020) (n=131) determined validity by comparing time allocations at a more granular (HETUS "two-digit") level than Harms et al. (2019), comprising 33 activity groups (sleep; eating; other personal care; main job; other paid work–related; school or university; food management; household upkeep; make, care for textiles; gardening and pet-care; construction and repairs; shopping and services; household management; childcare; organizational work; help to other households; participatory activities; social and entertainment; entertainment and culture; resting and time out; physical exercise; arts and hobbies; computing; games; reading; television; radio and recordings; work travel; education travel; unpaid work travel; civic travel; leisure travel; exercise travel). (48) The study found significant AWC-IAR versus TUD differences, using pairwise t-tests, differences for eating (53.7 vs 71.7 for AWC-IAR and TUD, respectively; p <.001), other personal care (57.6 vs 51.8 for AWC-IAR and TUD, respectively; p <.05), food management (52.7 vs 47.9 for AWC-IAR and TUD, respectively; p <.05), reading (29.9 vs 24.2 for AWC-

IAR and TUD, respectively; p <.05), and education travel (9.4 vs 11.8 for AWC-IAR and TUD, respectively; p <.05).

Laskaris et al. (2019) (n=51) determined validity by assessing AWC-IAR versus TUD agreement for work-related activities (burning; dismantling; sorting and loading; buying, selling, weighing; transporting materials; repair; other), non-work-related activities (sitting; smoking while sitting; eating or drinking while sitting; other), and transportation-related activities (walking; bicycling; motorbike or car). (69) Overall, agreement between the methods was found to be "none to slight" (kappa=0.17). No activity-specific kappa statistics were reported. Mean time allocation comparisons between AWC-IAR and TUD were also presented, however no statistical tests of significance were performed.

Discussion

For assessing time-use, there were three concurrent criterion validity studies of the AWC-IAR method, and none of the 24HR method. Two of the three AWC-IAR studies were separate analyses of the same UK-based data set, which considered the AWC-IAR method to be the reference rather than the comparator (TUD). Using AWC-IAR as a reference method is problematic given that, as revealed by this review, there is no evidence that it is a valid method for assessing time use. A strength of these two studies, however, is having used an international standard lexicon (the HETUS) of activity groups. For both studies, validity was assessed via comparisons of time allocations between the AWC-IAR and 24HR methods. Harms et al. (2019) analysed the results at a relatively high level (10 activity groups) and Gershuny et al. (2020) analysed the results at a relatively granular level (33 activity groups). (48,49) Both studies purported to analyse results at aggregate and individual levels, but key information was lacking, such as the statistical tests used and pertinent results. Nevertheless, consistencies across the studies (e.g., time allocated to eating being significantly higher when assessed by TUD than AWC-IAR) lend credibility to their respective findings. Both studies also reported the correlation between the time allocations derived via AWC-IAR and TUD, which is an inappropriate statistic for assessing validity between two methods. (109) Having used AWC-IAR as the reference method, no conclusions are made by the authors to its validity.

The AWC-IAR validation study conducted in Ghana had a small sample size (n=51), assessed time allocations to highly granular activities relevant to only a narrow field of interest (e-waste workers), and was also methodologically quite weak. Although mean time allocations via the AWC-IAR and TUD methods were presented, no statistical tests of significance were reported. Validity was assessed via an overall measure of agreement (kappa) and found to be quite poor. Nonetheless, the authors concluded that, based on this study, the AWC-IAR method would improve assessments of occupational exposures, which was not supported by the evidence presented.

CONCLUSION

Overall, this review indicates that there is insufficient evidence to determine the feasibility of AWC-based methods in rural, LIC contexts, or the validity of AWC-based or 24HR methods for assessing, specifically, dietary diversity or time use among mothers of infants and young children. The remaining gaps are substantial.

Regarding AWC feasibility, key participant-related challenges identified by studies conducted in UMICs and HICs in this review (e.g., physical and emotional burden, and difficulty operating the device) are likely to be even more prevalent in rural LIC contexts, where exposure to electronic wearable devices is limited. Moreover, rural LIC contexts likely pose additional challenges not addressed in these studies, such as widespread illiteracy, rugged living and working conditions, and uncertain cultural norms and weak policy environment regarding photography for research purposes in public spaces.

The number and extent of the validation analyses conducted thus far have also been limited. There are no studies validating the AWC-IAR method for assessing dietary diversity in any context, and no concurrent criterion validity studies of the AWC-IAR or 24HR methods for dietary diversity among infants and young children. There are no studies validating the 24HR method for assessing time use in any context, and no time-use studies at all that have targeted women, mothers of children under two years of age, or people living in rural contexts. There are also no AWC-IAR concurrent criterion validity studies conducted in any LIC.

Although most of the validation studies conducted statistical tests of comparison, no studies assessed other important aspects of validity, such as random error or the patterns in the error across the spectrum from lower to higher DDSs / time allocations as obtained via the Bland-Altman limits of agreement method. (109) Regarding time use, although Gershuny et al. (2020) describes the importance of accounting for simultaneous activities in calculating mean time allocations, none of the studies reported inter-method comparisons of the number of simultaneous activities performed by participants. (48)

Accurate and timely maternal and child dietary diversity and women's time-use data are needed to understand relationships between women’s time use and child/maternal dietary quality, and to identify populations at risk of undernutrition due to the consumption of poor-quality diets. AWCs have been used in HICs to prospectively collect dietary intake and time-use data and to ameliorate some of the data quality issues inherent to traditional recall methods. This review indicates, however, that more and more high-quality research is needed to determine if the AWC-IAR and 24HR methods are feasible, acceptable, and valid for assessing maternal and/or child dietary diversity and women's time use in rural LICs.

Table 2.1. Characteristics of included AWC feasibility studies

Author (year), country [LOD]	Device / Image capture interval	Topic / outcome	Population (sample size) / sampling method	Study duration	Key feasibility findings	Conclusion
Arab et al. (2011), USA [HIC]	bespoke / 10s	dietary intake / feasibility	healthy non-Hispanic Caucasian & African American adult men & women (i) (N=10) (ii) (N=14) / convenience	(i) 2-3 days (ii) 6-10 days	**Participant:** AWC was cumbersome to wear; Most reported no technical difficulties; Wearing the AWC was burdensome; IAR aided recall; IAR was not burdensome. **Community:** Not reported. **Researcher:** AWC inoperability; Reactivity; Image quality issue; Image processing not onerous.	a promising complementary technology for dietary recall, although challenges regarding the management of thousands of images and participant self-consciousness remain
Arab & Winter (2010), USA [HIC]	bespoke / 10s	dietary intake: frequency of imaging needed to capture all foods eaten	technologically savvy university employees (N=10) / convenience	1 day	**Participant:** No instances of non-compliance. **Community:** Not reported. **Researcher:** AWC inoperability; Reactivity; Image quality issue.	consumption of beverages and snacks are likely to be missed, even with AWCs having an image capture interval of 10s
Barr et al. (2014), New Zealand [HIC]	SenseCam / 10s	food environment: feasibility	children aged 12 years (N=6) / convenience	2 days	**Participant:** Wearing the AWC was not burdensome; Non-compliance reported. **Community:** Acceptability issue. **Researcher:** AWC inoperability; Reactivity; Image quality issue; Image processing onerous; Image coding onerous.	a feasible method for capturing children's exposure to food and beverage marketing
Beltran et al. (2018), USA [HIC]	eButton / 4s	dietary intake: feasibility and intercoder reliability	children aged 9-13 years (N=30) / convenience	2 days	**Participant:** AWC interfered with daily activities (excessive movement) and was physically uncomfortable to wear; withdrawal attributed to AWC. **Community:** Not reported. **Researcher:** AWC inoperability; Image quality issue; Image coding onerous.	feasibility challenges remain, which require additional training and/or technology improvement
Chambers et al. (2017), New Zealand [HIC]	Autographer / 7s	activity space: children's leisure time activity space	children aged 11-13 (N=168) / stratified	4 days	**Participant:** Not reported. **Community:** Not reported. **Researcher**: Data loss, reasons unknown; Image quality issue.	AWC-captured images, combined with GPS data, provides useful insight into children's activity spaces

Author (year), country [LOD]	**Device / Image capture interval**	**Topic / outcome**	**Population (sample size) / sampling method**	**Study duration**	**Key feasibility findings**	**Conclusion**
Cowburn et al. (2015), UK [HIC]	SenseCam / 10-15s	dietary intake & food environment: feasibility and acceptability	male and female students aged 13–15 years (N=22) / purposive	4 days	**Participant:** Most reported no technical difficulties; Wearing the AWC was not burdensome; refusal attributed to AWC; IAR aided recall. **Community:** Not reported. **Researcher:** AWC inoperability; Image quality issue; Image coding onerous.	AWC-captured images used in an IAR prompt detailed conversations and important insights into dietary behaviours, however feasibility challenges remain
Davies et al. (2020), Australia [HIC]	Autographer / 30s	physical activity: feasibility	adults aged 18–30-years (N=78) / purposive	3 days	**Participant:** Not reported. **Community:** Not reported. **Researcher**: Data loss, reasons unknown; Image quality issue; Image coding onerous.	AWC-captured images, combined with accelerometers, provides useful insight into physical activity behaviours, however AWCs are infeasible for large-scale studies
Doherty et al. (2013), New Zealand & USA [HIC]	SenseCam / 20s	physical activity: feasibility	university workers (N=52) / convenience	3 days	**Participant:** Difficulty operating the AWC. **Community:** Not reported. **Researcher:** Image quality issue.	the best available method for objectively categorising activity space
Everson et al. (2019), UK [HIC]	Autographer / 15s	physical activity: validity and acceptability	students aged 9–11-years (N=14) / purposive	2 days	**Participant:** AWC interfered with daily activities (excessive movement); Wearing the AWC was burdensome for some; refusal attributed to AWC; Non-compliance reported; Images deleted by participant. **Community:** Acceptability issue. **Researcher:** Reactivity; Image quality issue.	AWC-captured images, combined with accelerometers, provides useful insight into physical activity behaviours, however technical and ethical challenges remain
Gage et al. (2019), New Zealand [HIC]	Autographer / 7s	other (sun-protective behaviours): feasibility	children aged 11–13 years (N=15) / stratified	4 days	**Participant:** Not reported. **Community:** Not reported. **Researcher:** Image quality issue.	a feasible method for assessing sun-safety behaviour

Author (year), country [LOD]	Device / Image capture interval	Topic / outcome	Population (sample size) / sampling method	Study duration	Key feasibility findings	Conclusion
Gemming et al. (2013), UK [HIC]	SenseCam / 20-30s	dietary intake: validity	adults aged 18-65 years (N=13) / convenience	2 days	**Participant:** Wearing the AWC was burdensome for some; IAR aided recall. **Community:** Not reported. **Researcher:** AWC inoperability; Reactivity; Image quality issue.	AWC-captured images used in an IAR improves dietary recall
Gemming et al. (2015), New Zealand [HIC]	SenseCam / 20s	dietary intake: validity	adults aged 18–64 years (N=40) / convenience	4 days	**Participant:** Non-compliance reported; Decreased concern for content of AWC-captured images over time. **Community:** Not reported. **Researcher:** AWC inoperability; Image quality issue; Image coding onerous.	AWC-captured images used in an IAR improves dietary recall, although image capture intervals <20s may further improve accuracy
Gemming et al. (2015), New Zealand [HIC]	SenseCam / 20s	food environment: feasibility	healthy adults (N=40) / convenience	4 days	**Participant:** Non-compliance reported. **Community:** Not reported. **Researcher:** AWC inoperability; Image quality issue.	a feasible method for objectively capturing contextual information for dietary behaviours, as a complement to traditional recall methods
Gershuny et al. (2020), UK [HIC]	Autographer / 20-30s	time use: validity	adults, 18 years and older (N=148) / convenience	1 day	**Participant:** Non-compliance reported. **Community:** Acceptability issue. **Researcher:** Image quality issue.	No AWC-related conclusions reported.
Gurrin et al. (2013), Ireland [HIC]	bespoke / 6s	physical activity: feasibility	adults (N=47) / purposive	1 day	**Participant:** Non-compliance reported. **Community:** Not reported. **Researcher:** AWC inoperability.	Smartphones are a feasible alternative to stand-alone AWC devices for capturing health behaviour data
Harvey et al. (2016), UK [HIC]	SenseCam / 11s	physical activity: acceptability and feasibility	older adults (mean age: 68 years) (N=6) / convenience	7 days	**Participant:** AWC was not cumbersome to wear; Difficulty operating the AWC; Wearing the AWC was not burdensome. **Community:** No acceptability issue. **Researcher:** AWC inoperability; No reactivity; Image quality issue; Image coding onerous.	acceptability was high; however, feasibility was hampered by operational issues

Author (year), country [LOD]	Device / Image capture interval	Topic / outcome	Population (sample size) / sampling method	Study duration	Key feasibility findings	Conclusion
Imtiaz et al. (2020), USA [HIC]	bespoke / 1s	other (smoking): feasibility	smokers aged 19–70 years (N=10) / convenience	1 day	**Participant:** AWC was not cumbersome to wear; Wearing the AWC was not burdensome. **Community:** Not reported. **Researcher:** No reactivity; Image quality issue.	a feasible method for objectively capturing contextual information for smoking behaviours
Kamar et al. (2019), UK [HIC]	SenseCam / 20-30s	dietary intake & food environment: feasibility	children aged 11-16 years (N=8) / convenience	3 days	**Participant:** Poor AWC lanyard fit; Wearing the AWC was not burdensome; withdrawal attributed to AWC; IAR aided recall. **Community:** Acceptability issue. **Researcher:** AWC inoperability; Image quality issue.	a feasible method for assessing dietary behaviour among children
Kelly et al. (2011), UK [HIC]	SenseCam / 10s	physical activity & activity space: validity	well-educated adults aged 24-60 years (N=20) / convenience	<1 day	**Participant:** Forgot to wear the AWC. **Community:** No acceptability issue. **Researcher:** Image quality issue.	a feasible and acceptable method for assessing mode and duration of travel, which provides insight into the source of reporting errors in traditional methods
Kelly et al. (2014), UK, USA, New Zealand [HIC]	SenseCam / UNSP	activity space: validity and reliability	adults aged 19-60 years (N=84) / purposive	3-4 days	**Participant:** Forgot to wear and/or charge the AWC; refusal attributed to AWC. **Community:** Not reported. **Researcher:** AWC inoperability; Image quality issue.	provides insight into the magnitude of reporting errors in traditional methods
Kelly et al. (2012), UK [HIC]	SenseCam / UNSP	activity space: validity	adolescents aged 13–15 years (N=17) / convenience	UNSP	**Participant:** Forgot to wear the AWC; Difficulty operating the AWC; Wearing the AWC was burdensome for some. **Community:** Not reported. **Researcher:** Image quality issue.	a feasible method for assessing mode and duration of travel among children

Author (year), country [LOD]	**Device / Image capture interval**	**Topic / outcome**	**Population (sample size) / sampling method**	**Study duration**	**Key feasibility findings**	**Conclusion**
Kelly et al. (2015), UK [HIC]	Autographer / UNSP	time use: feasibility	adults aged 21-58 years (N==14) / convenience	1 day	**Participant:** AWC was cumbersome to wear (excessive movement); Wearing the AWC was not burdensome; Non-compliance reported; IAR aided recall; IAR was not burdensome. **Community:** Acceptability issue. **Researcher:** No reactivity; Image processing not onerous; Image coding onerous; High cost.	a feasible complementary technology, used in conjunction with a traditional recall method, for assessing time use
Kerr et al. (2013), USA [HIC]	SenseCam / 10-15s	physical activity: feasibility	cyclists aged 18–70 years (N=40) / convenience	3-5 days	**Participant:** No issues with the AWC; Wearing the AWC was not burdensome; No instances of non-compliance; Images deleted by participant. **Community:** Not reported. **Researcher:** Image quality issue; Image coding onerous.	a feasible complementary technology for assessing sedentary behaviour
Laskaris et al. (2019), Ghana [LMIC]	GoPro / 60s	time use: validity	male e-waste workers (N=142) / UNSP	<1 day	**Participant:** IAR reduced burden. **Community:** Not reported. **Researcher**: Data loss, reasons unknown; Image quality issue.	a feasible complementary technology for assessing occupational exposures
McKerchar et al. (2020), New Zealand [HIC]	UNSP / 7s	food environment: description of children's interactions with food in convenience stores	children aged 11–14 years (N=37) / simple random	4 days	**Participant:** Not reported. **Community:** Not reported. **Researcher:** Image quality issue.	No AWC-related conclusions reported.
Mckerchar et al. (2019), New Zealand [HIC]	UNSP / 7s	food environment: feasibility	children aged 11–13 years (N=37) / simple random	4 days	**Participant:** Not reported. **Community:** Not reported. **Researcher:** Image quality issue.	a promising method for assessing the food environment

Author (year), country [LOD]	Device / Image capture interval	Topic / outcome	Population (sample size) / sampling method	Study duration	Key feasibility findings	Conclusion
Nebeker et al. (2016), USA [HIC]	SenseCam / 20s	other (AWC ethics): acceptability	adult men and women, aged 18–85 years (N=82) / convenience	7 days	**Participant:** AWC interfered with daily activities; Wearing the AWC was burdensome for some; Images deleted by participant. **Community:** Acceptability issue. **Researcher:** Not reported.	concerns regarding physical discomfort and privacy among participants and their community warrant review of the AWC ethical framework
O'Loughlin et al. (2013), Ireland [HIC]	SenseCam / UNSP	dietary intake: validity	athletes aged 16-24 years (N=47) / convenience	1 day	**Participant:** Difficulty operating the AWC; IAR aided recall. **Community:** Not reported. **Researcher:** AWC inoperability; Image quality issue.	a promising complementary technology for assessing quantitative dietary intake
Oliver et al. (2013), New Zealand [HIC]	SenseCam / 10s	activity space: feasibility	adult university employees (N=15) / convenience	3 days	**Participant:** Not reported. **Community:** Not reported. **Researcher:** Image quality issue; Image processing not onerous; Image coding onerous.	a feasible method for assessing the built environment
Pearson et al. (2017), New Zealand [HIC]	Autographer / 7s	other (blue space exposure): description	children aged 11–13 years (N=166) / simple random	4 days	**Participant:** Forgot to charge the AWC; Non-compliance reported. **Community:** Not reported. **Researcher:** Image quality issue.	a reliable method for assessing blue space exposure
Raber et al. (2018), USA [HIC]	eButton / 4s	food preparation practices: feasibility	pre- and early adolescents aged 9-13 (N=13) / convenience	1 day	**Participant:** Not reported. **Community:** Not reported. **Researcher:** AWC inoperability; Image quality issue; Image coding onerous.	a feasible method for assessing food preparation behaviours among children
Raber et al. (2020), USA [HIC]	eButton / UNSP	food preparation practices: feasibility	parents of a child aged 5-17 years (N=40) / convenience	<1 day	**Participant:** Not reported. **Community:** Not reported. **Researcher:** Image quality issue.	a feasible complementary or reference method for assessing food preparation behaviours

Author (year), country [LOD]	Device / Image capture interval	Topic / outcome	Population (sample size) / sampling method	Study duration	Key feasibility findings	Conclusion
Schrempft et al. (2017), UK [HIC]	SenseCam / UNSP	food environment: feasibility	parents of children aged 2-8 years (N=15) / convenience	4 days	**Participant:** AWC interfered with daily activities; Forgot to wear and/or charge the AWC; Wearing the AWC was not burdensome; refusal attributed to AWC; Non-compliance reported. **Community:** Not reported. **Researcher:** Image quality issue; Image coding onerous.	a promising method for objectively assessing the food environment
Signal et al. (2017), New Zealand [HIC]	Autographer / 7s	food environment: description	children aged 11–13 years (N=168) / simple random	4 days	**Participant:** Wearing the AWC was not burdensome; Images deleted by participant. **Community:** Not reported. **Researcher:** AWC inoperability; Image quality issue.	a promising method for objectively assessing children's exposure to food and beverage marketing
Smith et al. (2019), New Zealand [HIC]	Brinno TLC / 15s	other (screen use): feasibility and acceptability	adolescents aged 13−17 years (N=15) / convenience	7 days	**Participant:** AWC interfered with daily activities (excessive movement); Forgot to power the AWC on / off; Non-compliance reported; Images deleted by participant. **Community:** Acceptability issue. **Researcher:** AWC inoperability; Image quality issue.	a feasible and acceptable method of assessing screen-time behaviours among adolescents
Smith et al. (2019), New Zealand [HIC]	Autographer / 7s	dietary intake & food environment: description	children aged 11-14 years (N=158) / simple random	4 days	**Participant:** Non-compliance reported. **Community:** Not reported. **Researcher:** Image quality issue.	No AWC-related conclusions reported.
Veatupu et al. (2019), Tonga [UMIC]	Autographer / 7s	dietary intake & food environment: description	children aged 10–12 years (N=36) / simple random	3 days	**Participant:** No refusals to participate; Non-compliance reported. **Community:** Not reported. **Researcher:** AWC inoperability; Image quality issue.	No AWC-related conclusions reported.
Watkins et al. (2019), New Zealand [HIC]	Autographer / 7s	other (marketing exposure): feasibility	students aged 11–12 years (N=12) / purposive	4 days	**Participant:** Not reported. **Community:** Not reported. **Researcher:** Image quality issue; Image coding onerous.	a feasible method for objectively assessing children's exposure to marketing

Author (year), country [LOD]	Device / Image capture interval	Topic / outcome	Population (sample size) / sampling method	Study duration	Key feasibility findings	Conclusion
Wilson et al. (2016), UK [HIC]	SenseCam / 30s	other (AWC acceptability): acceptability and feasibility	older adults with chronic pain (N=18) / purposive	7 days	**Participant:** AWC interfered with daily activities (excessive movement); Difficulty operating the AWC; Wearing the AWC was burdensome for some; refusal attributed to AWC; withdrawal attributed to AWC; Non-compliance reported; Images deleted by participant. **Community:** No acceptability issue. **Researcher:** No reactivity.	anticipated acceptability issues (at study launch) were a bigger challenge than experience acceptability issues, which contributed to non-compliance; operating the device was a feasibility challenge for this population of older adults
Zhou et al. (2019), China [UMIC]	Narrative Clip / UNSP	dietary intake: validity and feasibility	children aged 9–11 years (N=62) / purposive	7 days	**Participant:** AWC was not cumbersome to wear; No instances of non-compliance; IAR aided recall. **Community:** Not reported. **Researcher:** Not reported.	a feasible complementary technology for assessing dietary behaviour among children, especially effective for improving reporting of snack consumption
AWC, automated wearable camera; IAR, image-assisted recall; LOD level of development; HIC high-income country; UMIC upper-middle-income country; UNSP unspecified;						

Table 2.2 Characteristics of included dietary diversity criterion validity studies

Author (year), country [LOD]	**Comparator Method(s) / Reference Method (Validation Period)**	**Outcome(s) of Interest**	**Population, sample size / sampling method**	**Analyses (statistical tests): findings**	**Conclusion**
Hanley-Cook et al. (2020), Ethiopia [LIC] and Cambodia, Zambia [LMIC]	Retrospective: list-based and open 24HRs / Prospective: WFR (1 day)	MDD-W; DDS;	women of reproductive age, Cambodia (n = 430) / convenience, Ethiopia (n = 431) / representative, Zambia (n = 476) / convenience;	**Comparisons** **MDD-W** (McNemar's chi-square tests): Proportion achieving MDD-W was 30% (95% CI: 28–33) via WFR; 46% (95% CI: 43–48) via list-based 24HR; and 40% (95% CI: 37–42) via open 24HR (both $p < 0.001$). **DDS** (Wilcoxon matched-pairs signed-rank test): Median DDS was 4 (3,5) via WFR, list-based 24HR, and open 24HR, however the results for the 24HR methods were statistically higher than those of WFR (both $p < 0.001$). **Reliability** **MDD-W** (Cohen's kappa): Moderate agreement for the 24HR methods vs WFR (k = 0.51–0.57). **DDS** (Weighted Cohen's kappa): Moderate agreement for the 24HR methods vs WFR (k = 0.47–0.52). **Predictive power** M**DD-W** (receiver operator characteristic (ROC) analysis): Pooled AUCs for the list-based and open recall methods were significantly different from those of WFR (both $p < 0.001$).	Both list-based and open 24HR methods overestimate the prevalence of MDD-W and DDS in women of reproductive age in low- and middle-income countries
24HR 24-hour recall; DDS dietary diversity score; LIC low-income country; LMIC lower-middle-income country; LOD level of development; MDD-W minimum dietary diversity for women; WFR weighed food record.					

Table 2.3. Characteristics of included time use criterion validity studies

Author (year), country [LOD]	**Comparator Method(s) / Reference Method (Validation Period)**	**Outcome(s) of Interest** [activity groups]	**Population, sample size / sampling method**	**Analyses (statistical tests): findings**	**Conclusion**
Harms et al. (2019), UK [HIC]	Prospective: TUD / Retrospective: AWC-IAR* (1 day)	**time allocated to activities;** [HETUS single-digit activity groups: sleep, personal care; eating and drinking; paid work and related; unpaid work, childcare; voluntary/civic activity; social activity, relaxation; physical activity; games, hobbies; media: tv, radio, read, it; travel: all types;]	adults, 18 years and older (N=131) / convenience; snowball;	**<u>Comparisons</u>** **TA:** (no statistical tests of significance) With two exceptions (physical activity and eating), AWC-IAR vs TUD mean time allocations differed by <10%.	No AWC-related conclusions reported.
Laskaris et al. (2019), Ghana [LMIC]	Prospective: TUD / Retrospective: AWC-IAR* (1 day)	**time allocated to activities;** [work-related activities (burning; dismantling; sorting and loading; buy, sell, weigh; transporting materials; repair; other;); non-work-related activities (sitting; smoking while sitting; eating or drinking while sitting; other;); transportation-related activities (walking; bicycling; motorbike or car;)]	male e-waste workers aged 16-50 years (N=51) / UNSP	**<u>Comparisons</u>** **TA:** There appear to be substantial differences between TUD and AWC-IAR means for some activities, however no statistical tests of significance were reported. **<u>Agreement</u>** **TA:** (unweighted kappa) The agreement was low (0.17).	AWCs can improve occupational exposure assessments.
Gershuny et al. (2020), UK [HIC]	Prospective: TUD / Retrospective: AWC-IAR* (1 day)	**time allocated to activities;** [HETUS two-digit activity groups: sleep; eating; other personal care; main job; other paid work–related; school or university; food management; household upkeep; make, care for textiles; gardening and pet-care; construction and repairs; shopping and services; household management; childcare; organizational work; help to other households; participatory activities; social and entertainment; entertainment and culture; resting and time out; physical exercise; arts and hobbies; computing; games; reading; television; radio and recordings; work travel; education travel; unpaid work travel; civic travel; leisure travel; exercise travel;]	adults, 18 years and older (N=131) / convenience; snowball;	**<u>Comparisons</u>** **TA:** (pairwise t-tests) Out of the 33 activities, there were significant differences via AWC-IAR vs TUD in eating (53.7 vs 71.7, p-value UNSP), reading (29.9 vs 24.2, p-value UNSP), and watching television (64.1 vs 74.6, p-value UNSP).	No AWC-related conclusions reported.

AWC-IAR automated wearable camera-based image-assisted recall; HETUS, Harmonised European Time Use Surveys; HIC, high-income country; LMIC, lower-middle-income country; LOD, level of development; TUD, time-use diary; UNSP, unspecified.
* These studies consider AWC-IAR to be a reference method, despite that its validity has not yet been demonstrated.

Chapter 3: Methods

STUDY DESIGN

This study was nested within a cross-sectional study of women with a child aged between 12 and 23 months inclusive (n=211), to examine the impact of a labour-saving technology on the time women spent doing childcare, food preparation, feeding or serving others, and eating.[6] The study was conducted between January and February 2018 in Bugiri and Kamuli Districts, Eastern Region, Uganda.

As discussed in the Introduction (Chapter 1), in addition to evaluating the feasibility of the AWC-IAR method, this study assessed the concurrent criterion validity of both the AWC-IAR and 24HR methods, compared to observation for maternal and child dietary diversity, and women's time use.[7] "Criterion validity" refers to how

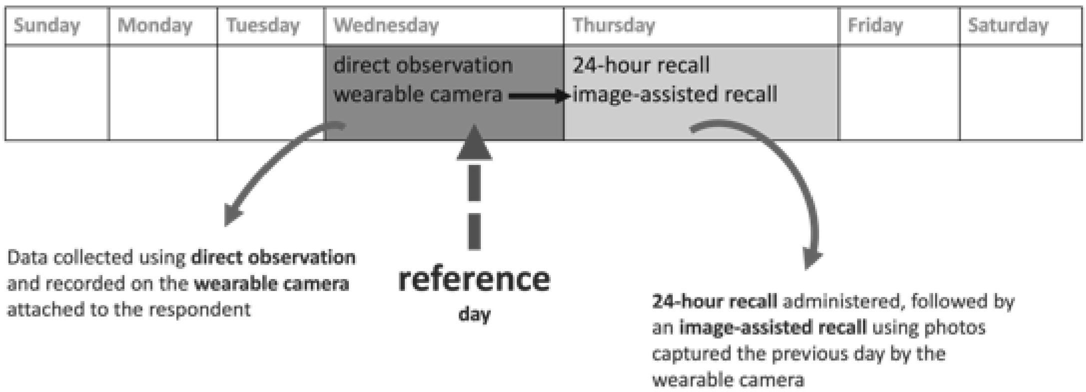
---

[6] The umbrella study, "Using Information Communication Technologies (ICTs) to understand the relationships between labour-saving agricultural innovations, women's time use and maternal and child nutrition outcomes", aimed to "enhance understanding of the impact of nutrition-sensitive agriculture interventions on women and young children in order to ensure they have a positive rather than a negative effect on maternal and child well-being". The specific objectives were to: (1) assess the feasibility of using an inter-active voice diary (later determined to be an interactive voice response [IVR] system accessed by participants via mobile phones) and a Global Positioning System [GPS] -linked wearable camera (later determined, due to cost constraints, to be two separate devices - an AWC and a GPS tracker) to assess women's workload, activity patterns and maternal and infant dietary intakes; (2) determine the relative validity of each of these two methods via 12-hour OBS; (3) compare the results generated using these two methods with those generated using traditional recall techniques; and (4) develop a framework of analysis for assessing the positive and negative impacts of alternative nutrition-sensitive interventions. Regarding objective (4), the study was designed to be conducted in 12 communities, half of which had access to a labour-saving device. The selected device was a maize sheller, which was being promoted by the Sasakawa Africa Association in these communities. However due to delays, data collection did not begin until after the post-harvest season, when the maize shellers would have been used. Only one out of 211 participants used a maize sheller during the period of data collection. Therefore, time-use comparisons could not be made between communities with and without a maize-sheller as originally intended.

[7] As part of the umbrella study, the concurrent criterion validity of the mobile phone based IVR was also evaluated for assessing maternal and child dietary diversity and women's time use. In addition to the AWC, participants also carried a mobile phone, receiving three IVR calls throughout the day, at 10:00 am, 3:00 pm, and 8:30 pm.

well the scores of the measurement instrument(s) agree with the scores on the gold standard. (110,111) That is, I compared results obtained (for dietary diversity and time use) from OBS to (a) results obtained (for dietary diversity and time use) via AWC-IAR and (b) results obtained (for dietary diversity and time use) via the traditional method, 24HR. "Concurrent" refers to scores generated by different tests at the same time. (111) In this study, the three methods were administered to the same population for the same reference day. (Figure 3.1)

Figure 3.1. Illustration of the reference day used in this concurrent criterion validity study



For dietary diversity, the indicators compared across methods were DDS and MDD, which were collected for mother and child. For women, DDS is measured on a scale of 0 to 10 food groups (0-10 food groups (grains, white roots and tubers, and plantains; pulses; nuts and seeds; dairy products; meat, poultry, and fish; eggs; dark green leafy vegetables; other vitamin A-rich fruits and vegetables; other vegetables; and other fruits). (112) The MDD threshold for women is five out of the ten food groups. For children, DDS is on a scale of 0 to 7 food groups (grains, roots, and tubers; legumes and nuts; dairy products; flesh foods; eggs; vitamin A-rich fruits and vegetables; and other fruits and vegetables). (113) The MDD threshold for children is four out of the seven food groups. These analyses were conducted prior to the launch of the updated guidelines for administering the DDS for both women and children in 2021. (9,101)

For time-use, the indicators compared across methods were the number of minutes allocated to 42 activities, categorized into nine mutually exclusive International Classification of Activities for Time-Use Statistics (ICATUS) activity groups (employment; production for own use; domestic services; caregiving, unpaid volunteer; trainee and other unpaid work; learning; socializing and leisure; and self-care) and the number of concurrent activities performed across all 15-minute timeslots.

Data were collected over five consecutive days, following one of two possible patterns. (Table 3.1) Specifically, for both patterns, on day 1, eligibility was confirmed, a structured socio-demographic and health questionnaire ("mother's questionnaire") was administered, and anthropometric data were collected for all participants. For half of the study population, on day 2, maternal and child food consumption and maternal time allocation data were collected using OBS and recorded on the AWC attached to the respondent.

On day 3, the enumerator annotated the foods she thought were consumed and the activities she thought were undertaken by the participant, using the IAR form, based on what she could see in the photos captured on day 2 by the AWC (i.e., the enumerator image interpretation). The enumerator demarcated the series of foods and activities for review later that day with the respondent. Upon meeting with the respondent, maternal and child dietary diversity and women's time use were estimated using a 24HR method, followed by an IAR using the AWC-captured images. The enumerator revised her original annotations of foods / beverages consumed and activities undertaken by the respondent, as needed, based on the respondent’s feedback. On day 4, time allocation data were again recorded via AWC only (i.e., no observation). On day 5, an IAR was administered to estimate maternal and child dietary diversity and maternal time use using photos captured on day 4 by the AWC.

The other half of the study population began with the AWC only (i.e., days 2 and 4 were switched) and ended with all three methods (i.e., days 3 and 5 were switched). For all participants, on the 5^{th} day, a final structured questionnaire (the "technology questionnaire") was also administered to capture information on their experience using the different methods. Data collection was distributed across all days of the week at the population level to account for a day-of-the-week effect, and for each respondent, the enumerator assigned to conduct the OBS was different from the enumerator assigned to administer the 24HR and IAR.

Table 3.1. Household data collection patterns

Pattern	DAY 1	DAY 2	DAY 3	DAY 4	DAY 5
1	ICF MQ ANTHRO	OBS AWC+IM	24HR EII & IAR	AWC+IM	EII & IAR TQ
2	ICF MQ ANTHRO	AWC+IM	EII & IAR	OBS AWC+IM	24HR EII & IAR TQ

ICF, informed consent form; MQ, maternal questionnaire; ANTHRO, anthropometry; OBS, direct observation; AWC, automated wearable camera; IM, innovative methods questionnaire; 24HR, 24-hour recall; EII, enumerator image interpretation; IAR, image-assisted recall; TQ, technology questionnaire.

PARTICIPANTS AND SAMPLING

Twenty-two villages were purposefully selected from Bugiri and Kamuli districts of the Eastern Region, for this study, of which eleven had access to labour-saving technology and eleven did not. These villages participated in the Sasakawa Global 2000 Uganda (SG2000 Uganda) country programme (the local implementing partner for the parent study).

The sample size calculation (n=264; twenty-two communities, twelve households per community) was based on requirements of the time allocation study within which this current study was nested. A sample size of 132 participants per group enabled detection of a 30-min inter-group difference between women with access to a labour-saving device and households without access to a labour-saving device, assuming a standard deviation (SD) of 49 min, a design effect of 1·47, 80% power and P value of 0·05 and allowing for 10% attrition.

Data for time saved via labour-saving devices and design effect is sparse. The SD used here originated from Ounpuu (1988). (114) Due to the lack of pertinent reference studies, a statistician (Stephen Young, NRI) simulated the design effect.[8] Regardless, the design effect was not critical for this validation study because comparisons were made between results derived from different methods within the same population rather than between different populations. The sample size for the

[8] The design effect for this study was assessed retrospectively and found to be lower (1.04 for child dietary diversity) than the 1.47 used.

parent study was more than adequate for the Bland–Altman method of analysis. (109,111,115,116)

The sampling frame, for each village, was a household listing of all mothers with children born between 1 January 2016 and 1 May 2017 inclusive (to recruit children aged 12–24 months at the time of data collection). Sasakawa community-based facilitators visited households in their communities to collect the names of the household head, mother and child, and child's age, for those who qualified. Twelve mother–child dyads in each village were randomly selected to participate in the study; substitutions were made, as needed, until twelve mother–child dyads who met the inclusion/exclusion criteria were recruited.

Mother–child dyads were excluded if the child was <12 months or >23 months of age, was not yet eating solid foods on a regular basis, was a multiple-birth child, the mother was unable to communicate in Lusoga, Luganda or English; either the mother or child had a severe disability; the mother was not the biological mother of the child; the mother was a co-wife with another mother selected to participate in the study; or either the mother or child was not available for the duration of the study.

RECRUITMENT AND SENSITIZATION

The Sasakawa Africa Association was instrumental in accessing the study communities, garnering the trust of local leaders and community members, and completing the study in the allotted six-week timeframe. Following random selection of mother-child dyads, Sasakawa community-based facilitators notified mothers of their selection and invited them and their families to participate in a village-level sensitization meeting. Selected mothers were encouraged to bring their husband or other household or family members who were important for determining whether they could participate in the study. Mothers were also asked to bring their child’s immunization card or other official document so that the study team could confirm the child's age.

Sensitization and recruitment were led by one of the study principal investigators (PIs) (Joweria Nambooze), with support from local leaders and accompanied by a Sasakawa community-based facilitator. During the sensitization meeting, selected mothers and members of the community were provided an overview of the study objectives and activities, key aspects of informed consent and incentives, and each of the methods that would be administered. All equipment used in the study, including the devices such as the AWC, were demonstrated. The meeting concluded with a question and answer session and summary of key points. The sensitization meeting took approximately two-and-a-half hours. See Appendix 2 for the sensitization agenda.

Participating mothers were given a bar of soap, one kilogram of sugar, a half-litre of cooking oil and a T-shirt on the final day of the study. On the final day, they were also given a photo of their family taken by a supervisor using a polaroid camera.

PROTOCOLS AND INSTRUMENTS

As noted in Chapter 1, this thesis was part of a broader research study to assess the validity of three ICT methods: an IVR system using rudimentary (button) mobile phones, GPS trackers and the AWC. The focus of the thesis is the AWC combined with an IAR (AWC-IAR). Some details about the other methods are included below in footnotes, however, for context.

Structured questionnaires

The enumerators administered two structured questionnaires to the mother.

The first questionnaire collected information on: household socio-demographics, wealth (adapted from the Uganda 2012 PPI), expenditure and production (adapted from the Abbreviated WEAI); knowledge, attitudes and practices regarding infant and young child feeding and care; factors related to women's empowerment (adapted from the Abbreviated WEAI) and the child's health (adapted from the DHS). This paper-based questionnaire was administered on the first day of the data collection period, immediately following the informed consent. See Appendix 3 for the mother's questionnaire.

The second questionnaire was administered at the end of each participant's period of data collection (i.e., on day 5). Response options in this questionnaire included a four-point Likert for questions related to participants' perceptions of their experiences with each of the four data collection methods (AWC-IAR, 24HR, OBS, IVR), and categorical scales for questions related to mobile phone access, ownership and use, and willingness to participate in a future study using the 24HR, AWC-IAR or IVR, in addition to open-ended comments on their experiences. This questionnaire was kept brief, in consideration that it came at the end of five days of intensive data collection. Ideally, a less structured method, such as a focus group discussion, would have been used to elicit more in-depth feedback on the feasibility and acceptability of the AWCs, however we felt that the burden on participants and enumerators was too high. See Appendix 4 for the technology questionnaire.

A third brief semi-structured questionnaire (the "innovative methods" questionnaire), administered at the end of each AWC data collection day, is described below.

Anthropometric measurements

On the first day of data collection, following administration of the questionnaire, duplicate, serial anthropometric measurements of weight and height/length were taken from participating mothers and children. The weight of the mother and child were measured using a SECA digital floor scale with mother-baby tare (± 100 g), and the mother's height and child's length were measured using a microtoise wall-mounted measuring tape (± 1 mm / 1/8") and UNICEF length board (± 0.2 cm), respectively. At the start of each anthropometric data collection day, the SECA scale was calibrated using an object of known weight, and the length board and microtoise were calibrated using a rod of known length. Microtoise adjustments (the distance from the end of the tape at its longest point to the floor) were recorded on the anthropometry form, as needed. An enumerator and supervisor experienced in anthropometry did the measurements. One positioned the participant and the other made and recorded the measurement. See Appendix 5 for the anthropometric measurements form for mothers and children.

Weighed food records

A one-day WFR was the criterion diet assessment method. Specifically, enumerators weighed and recorded all foods/beverages consumed by mothers and children, from approximately 06.00 to 21.00 hours using dietary scales (±1 g, Salter Disc Electronic Digital Scale Model 1036) and a standard WFR protocol. (117) Recipe data were also collected by weighing the recipe ingredients and final cooked food and recording the cooking methods (e.g., fried, boiled and stewed). If the child was left in the care of another person, the enumerator remained with the mother and information about foods or beverages consumed, while the child was away, were collected from the secondary caregiver via recall upon their return. The amounts of foods/beverages consumed by the mother or child before 06.00 or after 21.00 hours were recalled and recorded. See Appendix 6 for the weighed food records instruments and related guides.

Direct time-use observations

For the criterion time allocation assessment method (i.e., OBS), enumerators recorded all activities undertaken by the respondent in 15-minute intervals ('timeslots') from approximately 06:00 to 21:00, using a structured paper-based instrument comprising 44 activities. The time-use matrix used to record activities was based on a module developed for the WEAI, which was itself based on the Lesotho Time Budget Study. (103,118) See Appendix 7 for the OBS instruments.

24h dietary recalls

On the day after the WFR was collected, two semi-quantitative multiple-pass 24h dietary recalls were administered to the mother to collect information on all foods and beverages consumed the previous day – one for her child and then one for herself. (117) For each recall, in the first pass, the mother was asked to list everything she (or her child) consumed the previous day; in the second pass, additional details about each food were recorded, including the time of consumption and ingredients in mixed dishes. In the third pass, mothers were asked to confirm the food groups consumed and not consumed. Specifically, for themselves and their child, mothers were queried regarding thirteen food groups, using a list-based method adapted for Uganda, that could be condensed into the 10- and 7-level dietary diversity score indicators used for women and children 6-23 months of age, respectively[9]. (112,113) The enumerator first listed all of the food groups the mother reported as having been consumed the previous day and verified, "is that correct?" The enumerator then listed all the food groups the mother reported as having not been consumed the previous day and verified, "is that correct?" Any discrepancies were resolved and recorded in a separate column. The quantity of each food consumed was not recorded. See Appendix 8 for the 24h dietary recall instruments.

---

[9] The thirteen food groups comprised: (1) foods made from grains, (2) white roots and tubers and plantains, (3) vitamin A-rich vegetables, roots and tubers, (4) dark green leafy vegetables, (5) other vegetables, (6) vitamin A-rich fruits, (7) other fruits, (8) organ meat, meat and poultry, and fish and seafood, (9) eggs, (10) pulses (beans, peas, and lentils), (11) nuts and seeds, (12) milk and milk products, and (13) insects and other small protein foods.

24h women's activity recalls

Also, on the day after the observation day, a multiple pass 24h activity recall was administered to the mother to collect information on all activities she performed the previous day. The protocol followed the same approach used in the 24h dietary recall. In the first pass, the respondent was asked to list everything she did the previous day; in the second pass, additional details about each activity were probed to help the enumerator classify it into one of the 44 activity categories and any concurrent activities were recorded. The time and duration of each activity performed was recorded in 15-minute increments using a structured paper-based instrument, similar to that used for OBS. In the third pass, the enumerator confirmed with the respondent that each activity was recorded correctly. See Appendix 9 for the 24h women's activity recall instruments.

On average, the three 24HRs (i.e., the 24h dietary recall for the child, and the 24h dietary recall and 24h activity recall for the mother) took approximately 72 minutes to complete.

AWC

The AWC used in this study was the iON SnapCam Lite (dimensions 42 x 42 x 13 mm)[10], a small and lightweight wearable camera that automatically took a picture every thirty seconds, which was saved to a memory card. (Figure 3.2) Examples of the photos obtained by device are provided in Figure 3.3.

[10] To achieve the sample size required in the allotted timeframe, more than thirty AWCs were needed. It was critical that all participants used the same AWC model to minimize variability. The device was therefore selected based on both cost and availability. At the time of purchase, the iON SnapCam Lite was no longer manufactured but could still be purchased from resellers for less than $100.

Figure 3.2. The iON SnapCam Lite

Figure 3.3. Illustrative photos obtained from the iON SnapCam Lite AWC

The AWC was attached to a study-provided t-shirt worn by the participant at approximately 06:00 and removed at approximately 21:00. (Figure 3.4) For administrative purposes, at the start of each data collection day, the enumerator was responsible for taking a single picture of the participant, her child and a placard displaying her study ID, using the designated function of the AWC. The enumerator reminded each participant of key points covered during sensitization, i.e., that at any time during the study she could remove or cover the device or request all images to be deleted; that the device was splash-proof but should not immersed in water; and to do exactly the activities she would have normally done. The AWC automatically recorded a picture every 30-seconds, storing all photos (approximately 1,500) on a memory card.

Figure 3.4. The AWC affixed via a bespoke clip to the neckline of participants' clothing

At the end of the data collection day, the AWCs were collected from participants by their assigned enumerator, aggregated by each team supervisor, and returned to me for processing. Upon receipt of the devices, I saved a copy of the images recorded on each AWC memory card to a local drive. I then assigned a tablet (16GB Samsung with a 10” screen, using Simple Gallery software for image display) and two participants' memory cards (still containing the 1,500 images) to an enumerator having no prior engagement with the participant (e.g., OBS or 24HR).

Semi-Structured "innovative methods" questionnaire

While retrieving the AWC from the study participant at the end of the data collection day, a brief "innovative methods" questionnaire was administered to assess the participant's experiences wearing the AWC, including any technical issues or reactions from members of their household or community. Each participant was also asked to reconfirm her consent to use the images captured by the AWC. See Appendix 10 for the "innovative methods" questionnaire.

AWC-based enumerator image interpretation

The next day, in the morning, each enumerator inserted the memory card for the first assigned participant into the tablet to review the photos captured by the AWC. In the left-most column of the IAR instrument (i.e., paper-based form), (See Appendix 11) the enumerator annotated the foods / beverages she thought were consumed and the activities she thought were undertaken by the participant, and their corresponding image numbers, based on what she saw in the photos i.e., the enumerator image interpretation. Based on her interpretation, the enumerator demarcated the series of foods and activities for review later that day with the respondent. After completing the annotation for the first assigned participant, the enumerator completed the same steps for the second assigned participant. Each enumerator image interpretation took approximately 1-2 hours to complete.

AWC-based image-assisted recall

Upon meeting with the respondent in the afternoon (i.e., the same day, after completing the enumerator image interpretation), the enumerator administered the 24h dietary recall for the child, followed by the 24h dietary recall and 24h activity recall for the mother, i.e., three 24HRs. The enumerator then oriented the mother to the first-person perspective photos by viewing, on the tablet (16GB Samsung with a 10” screen, using Simple Gallery software for image display) with the inserted memory card, five pre-selected photos captured by the AWC: a picture of the mother herself, a picture of her child, a picture of her home, a picture of her garden, and a picture where her own hand is visible while performing a task (e.g. while preparing or cooking food, digging, or using a mobile phone). The enumerator rated and recorded the participant's ability to recognize the content of these five photos on a three-point scale: recognized, recognized with help, or did not recognize.

The enumerator then administered the IAR. During this interview, the enumerator used "verbal probing" to elicit from the participant additional relevant information about the activities performed, for example to elaborate on what she was doing, who she was with, where she was going and why, etc. (119,120) Now using the right-most column of the IAR instrument, the enumerator revised her original annotations (i.e., the enumerator image interpretation) of foods/beverages consumed and activities undertaken by the respondent, as needed, based on the respondent's feedback. Unlike the 24HR method, in the IAR, information about the foods/beverages consumed and activities performed were elicited simultaneously, in the order in which they occurred during the previous day. On average, each IAR took approximately 78 minutes to complete.

The IAR protocol was adapted from one described by Kelly et al. (2015). (47) The protocol followed ethical guidelines for AWC research to ensure privacy of the participants was maintained. (106) Participants were free to request deletion of any or all photos during the IAR. See Appendix 11 for the IAR instruments.

Field notes

Each ICT used in this study (AWCs, GPS trackers, mobile phones) were labelled with a unique ID for the purposes of logistics and quality control. Over the six weeks of data collection, I recorded the AWCs found to be inoperable once deployed in the field using this ID. I also recorded in field notes other issues reported by the data collection team that may have affected study participation or compliance. For example, in Kamuli, fear among study participants periodically arose of the AWCs being a nefarious tool of recruitment into "the Illuminati". I noted these instances in order to be able to later assess whether the fear of "the Illuminati" influenced participants to be non-compliant or drop-out of the study.

PILOT TESTING

All instruments and protocols were pilot tested prior to the start of data collection, in rural areas of Central Region, Uganda. Special emphasis was given to the AWC and AWC-IAR protocols, given that they were truly novel. These tools underwent iterative piloting (approximately six iterations between late 29 November and 28 December 2017), with modifications to the instruments and protocols after each round.

FIELD TEAM TRAINING

The study field team (enumerators and supervisors) were recruited by a study PI (Joweria Nambooze). All members of the field team had a university degree and were fluent in in Luganda and/or Lusoga. Enumerators had prior experience collecting nutrition data. To avoid potential gender-based sensitivities posed by having an observer at the home for an extended period, all enumerators were female. Supervisors had prior experience leading teams and collecting nutrition and/or other sensitive data among low-income Ugandan households. Training for all protocols and instruments took place over one week (December 18-22, 2017). The training comprised classroom training, role-play practice, and an assessment with individualized feedback. Training also included two days of field practice.

DATA COLLECTION AND MANAGEMENT

Aside from the images captured by the AWCs, all data were collected on paper, due to the tight timeline between instrument piloting and refinement and data collection. The forms and equipment requirements varied according to the data collection day. (See Table 3.1. Household data collection patterns.) I made household assignments, prepared all forms, and oversaw outgoing and incoming materials (new and completed forms, all WFR equipment and AWCs) at the start and end of each data collection day.

At the end of each data collection day, before leaving the field, the supervisors reviewed and signed off on all completed forms collected from their team of enumerators. Upon returning from the field, the supervisors turned in the completed and reviewed forms and all equipment to me. I checked all forms and equipment to ensure that no data or supplies had been left behind. The completed forms were organized and stored in a locked trunk. I then downloaded the data from each AWC (and other devices) to a secured laptop while charging the devices and tablets in preparation for the next day. After the devices were fully charged, I compiled the pre-assigned packets of instruments and equipment required for the next day by supervisor.

As most of this work occurred at night, from about 9pm to 3am, Gwen Varley (another PhD candidate) was responsible for distributing the pre-assigned and collated materials (clean data collection forms and newly charged AWCs) to the supervisors at the start of the following day, around 5am.

During the day, while the field teams were collecting data, I spot-checked the completed WFR, OBS, 24HR, and IAR forms. In particular, during the first two weeks of data collection, I paid special attention to the quality of the WFR and IAR data. During the first week in the field, Elaine Ferguson provided supplementary WFR training. During the second week in the field, I provided personalized feedback to enumerators and additional training to those who were still struggling with these intensive methods.

Gwen Varley performed quality control for the completed mother's questionnaires.

See the checklists for the coordinator (Appendices 12), supervisors (Appendix 13) and enumerators (Appendix 14).

DATA ENTRY

In general, qualitative data were entered via tools I created in EpiData, whereas more structured data (e.g., the time-use matrices) were more easily entered by the RAs via Excel.

Administrative, anthropometric, and socio-economic data

I entered the IMMANA administrative and anthropometric data into Excel from the household listings, informed consent form, recruitment form, and anthropometry form. The socio-economic data (from the mother's questionnaire) was entered and cleaned by Gwen Varley and Joweria Nambooze.

Dietary diversity data

Two RAs (Nakimuli Sarah and Bernice D. Galinda) double-entered the dietary diversity data using EpiData forms that I created. WFRs were entered separately for mothers and children using two EpiData forms: (1) food intake and (2) recipes. In order to expedite analysis, WFR data (from food intake and recipes forms) were used to double-enter diet diversity data directly (instead of deriving it from the WFR) using an EpiData form. One WFR-based dietary diversity form was completed for each mother and one for each child.

For the 24HR dietary diversity data, one EpiData form was completed for the data collected from the mother and one for the child. Both the mother's initial report for each group and her final report, after having been asked to confirm the foods / beverages consumed and not consumed (as described above) were entered, so that changes between the initial report and confirmation could be assessed. Pass 1 (open recall of all foods consumed) and pass 2 (details for each food) data were not entered, due to limitations in available resources.

Similarly, for the IAR dietary diversity data, one EpiData form was completed for the mother and one for the child.

Time-use data

The OBS and 24HR time-use data were double-entered by Sarah and Bernice into an Excel-based matrix with each column representing a 15-minute time slot and each row representing a time-use activity. One worksheet was completed for each participant day. For ease of data entry, the Excel template had the same format and structure as the paper-based data collection form. I imported the Excel worksheets into Stata, appended and reshaped the data into one database with all participants and timeslots for OBS and one database with all participants and timeslots for 24HR. For the 24HR, pass 1 (open recall of all activities) and pass 2 (details for each activity) data were not entered, due to limitations in available resources.

The IAR time-use data, collected via an open- rather than closed-format, were double-entered by Sarah and Bernice. At the start of data entry, the RAs numbered each series of images on the completed IAR forms. Using an EpiData form that I created, the RAs entered the administrative data (household ID, enumerator ID, visit number, date of record, start time, end time, etc.) for each completed form once. They then entered data (series number, starting image number, ending image number, activities performed, etc.) for each series of images, including the activities performed. The activities performed per each series were recorded in two ways. First, the RAs typed the activity description verbatim, as it was recorded by the enumerators in the field. Second, the RAs coded the activity description into the 44 standard IMMANA activities (i.e., the same activities used in the OBS and 24HR data collection forms). Each series was coded with one or more activities. I imported the completed EpiData forms into Stata, appended and reshaped the data into a single database with all participants and their respective image series.

The RAs therefore did not interpret the AWC images directly, but rather coded the image series using the descriptions recorded in the field. I made a rubric to handle instances where insufficient information was provided by the field enumerator to conclusively determine the activity, so the data would be coded consistently. For example, mothers could have been chopping and peeling cassava or shelling groundnuts for either immediate consumption (coded as food preparation) or for storage (coded as post-harvest processing). Where insufficient information was provided, these activities were coded as food preparation.

Technology questionnaire and AWC data

Data from the technology questionnaire and innovative methods questionnaire were also double entered by Sarah and Bernice into an EpiData form that I created. I reviewed all image sets captured by the AWC to ensure that they were correctly identified per the date and ID displayed on the nameplate at the start of each data collection day.

DATA CLEANING

I imported all data into Stata for verification, cleaning, processing and use in analyses. For double-entered data, I verified discrepancies against the paper forms, and made changes as needed.

DATA PROCESSING

Anthropometric and socio-economic data

Weight-for-age z-score, length-for-age z-score and weight-for- length z-score were calculated for each child using the 2006 WHO growth standards, and BMI was calculated for each mother by dividing her body weight by her height squared. (121) To protect privacy, all mothers were weighed regardless of their pregnancy status, although women who had indicated in the first questionnaire that they were pregnant were excluded from anthropometric analyses. The proportions of children who were underweight (<–2 SD from median weight-for-age z-score), stunted (<–2 SD from median length-for-age z-score) and wasted (<–2 SD from median weight-for-length z-score) were calculated.

The proportions of mothers who were thin (BMI < 18·5 kg/m2), normal (BMI 18·5–24·9 kg/m2) and overweight/obese (BMI ≥ 25·0 kg/m2) were also calculated. The Ugandan 2012 PPI was calculated, as well as the proportion of the population living below $1·25/d. (122)

Dietary diversity data

The foods/beverages recorded over a period of 24 hours were coded into food groups, and a DDS was calculated for each mother and child for each method (WFR, 24HR and AWC-IAR). The DDS for children was based on seven food groups used to calculate the minimum dietary diversity indicator for infant and young children:

namely grains, roots and tubers; legumes and nuts; dairy products; flesh foods; eggs; vitamin A-rich fruits and vegetables; and other fruits and vegetables. (113) The DDS for women was based on the ten food groups used to calculate the minimum dietary diversity for women indicator: namely grains, white roots and tubers, and plantains; pulses; nuts and seeds; dairy products; meat, poultry and fish; eggs; dark green leafy vegetables; other vitamin A-rich fruits and vegetables; other vegetables; other fruits. (112,123) The percentages of children and women achieving minimum dietary diversity (MDD) was calculated using the 2010 and 2016 guidelines, respectively because data collection for this thesis preceded the recent updates to the WHO-UNICEF (9) and FAO (101) guidance. Therefore, the threshold MDD used for women was five food groups out of ten, and the threshold MDD for children was four food groups out of seven.[11]

Time-use data

Additional processing was required to convert the IAR time-use data to the structure used by OBS and 24HR for comparison. As elaborated in Chapter 4, the AWC's timestamp meta-data was unreliable, seemingly due to a flaw in the device. As a work-around, I used the start time on the innovative methods form and the fact that the AWCs captured one photo every 30 seconds to identify the starting time and ending time for each image series and the activities the series comprised. The derived start and end times per activity were used to generate a time-use matrix with each column representing a 15-minute time slot and each row representing a time-use activity for each respondent day. The IAR time-use data was then reshaped into a single database with all participants and their respective image series (i.e., the same structure OBS and 24HR).

Although time allocations to discreet activities may be useful for some research purposes, for the purpose of method validation, it is more practical to compare time allocations to activities at a higher level. We had initially intended to analyse the results based on the activity groups used in the WEAI time-use module, however in 2019 (after data collection but before analysis) the United Nations Statistics Division

[11] Per the 2010 WHO-UNICEF guidelines, breast milk intake was not included in the DDS calculations.

launched the ICATUS-2016 framework in an effort to standardize time-use statistics for the purpose of internationally comparison. (103,124) The advantages of using the ICATUS framework included international comparability, and also the ability to analyse the data at six levels of granularity.

For all methods, then, the number of minutes allocated to each of 44 activities recorded over the fifteen-hour period was calculated for each respondent and for each of the 3 data collection methods, in 15-minute intervals ('timeslots'). The discrete activities were categorized into the nine mutually exclusive International Classification of Activities for Time-Use Statistics (ICATUS) major divisions ("activity groups"): (1) employment and related activities ("employment"), (2) production of goods for own final use ("own production"), (3) unpaid domestic services for household and family members ("domestic chores"), (4) unpaid caregiving services for household and family members ("caregiving"), (5) unpaid volunteer, trainee and other unpaid work, (6) learning, (7) socializing and communication, community participation and religious practice ("socializing"), (8) culture, leisure, mass media and sports practices ("leisure"), and (9) self-care and maintenance ("self-care"). (124) See Appendix 15 for the ICATUS time-use activities and corresponding major divisions.

Of the 44 activities tracked in this study, four were considered to be "simultaneous", i.e., they could be performed while also performing other activities: care of the index child, care of other children or adults, chatting with friends or family, and watching TV or listening to the radio.[12] When just one activity was performed in a timeslot,

---

[12] We explore the implications of several options for processing the time-use data on total time allocation, most important being the traditional approach of equally proportioning time across all activities. Although this approach keeps "real time", in the sense that a day never exceeds 24 hours, it treats simultaneous activities as if they were instead performed sequentially. I assert that this is neither an accurate nor fair reflection of mothers' time allocated to care of infant children, at least not in this rural Ugandan context. When we made this decision, I believed that, for the purposes of method validation, the choice between alternative approaches to time-use processing was less important than it would be in a typical time use study, given that the time use data was analysed the same way across all three methods. This assumption was born out for the AWC-IAR validation, because we found that the median number of concurrent activities identified was the same via AWC-IAR and OBS. For 24HR, however, the median number of concurrent activities identified was

the activity performed counted for the entire 15-minutes. The simultaneous activities were always credited the full 15 minutes. However, for all other activities, when more than one activity was performed per timeslot, the 15 minutes were evenly distributed across the activities performed. For example, if in a 15-minute timeslot, the participant was caring for livestock (own production) and then started preparing food (domestic chores) while caring for the index child (caregiving), then caregiving–a simultaneous activity–was credited 15 minutes and own production and domestic chores were each credited 7.5 minutes.

DATA ANALYSIS

Data were analysed using Stata/SE version 15.1. P values <0·05 were considered significant for all tests. Details regarding each analysis (the feasibility and the validations for assessing maternal and child dietary diversity and women's time use) are included in each chapter's respective data analysis section. The dietary diversity and time-use validations followed similar data analysis processes. I will provide an overview of the analyses and rational for methods used here.

As noted above, for dietary diversity, the primary outcome variables analysed for both mothers and children were DDS and MDD. Following convention, the ordinal DDS variable (ranging from 0 to 7 for children and 0 to 10 for women) was treated as a continuous measure. Cases with incomplete maternal and/or child dietary diversity data for any of the three methods (WFR, 24HR or AWC-IAR) were eliminated from analysis.

Although the intention was to assess dietary diversity over 15 hours, actual observation starting and ending times differed between participants. Dietary intake data outside of these observed hours would have been based solely on participant recall. Only a third of participant days exceeded 14 hours of observation whereas approximately 80 percent of participant days exceeded 13 hours of observation. Therefore, a minimum threshold of 13 hours of observation and photos from the

---

significantly lower than was identified via OBS, opening the possibility that the choice of time-use processing approach may have affected the validation results. As noted in Chapter 6, the trade-offs and implications of different time-use processing approaches warrant further exploration and transparency.

AWC were deemed adequate to ensure the integrity of the inter-method comparisons. Any cases not meeting this threshold were eliminated from analysis. Out of 422 participant days (2 days per participant), 92 (22%) were excluded from analysis due to having less than 13 hours of AWC-captured images.

As also noted above, for time use, the primary outcome variables analysed were the total number of minutes allocated to each of the nine ICATUS major divisions (minutes/day) and the number of concurrent activities performed across all 15-minute timeslots. Due to the inter-participant differences in actual observation start and end time and technical challenges with insufficient light in the early morning and evening, the analyses were limited to the 12-hour period from 8am to 8pm. Out of 422 participant days, 13 (3%) were excluded from analysis due to having less than 12 hours of AWC-captured images.

Comparisons of key socio-demographic characteristics for participating and missing households

First, I compared socio-demographic characteristics for women and children included in the study versus those who dropped out or were excluded from the study. Key socio-demographic characteristics for participating and missing households were compared using the Mann–Whitney[13] two-sample statistic for continuous data, and the Fisher Exact[14] test for categorical data. The Fisher Exact test was used for the categorical data instead of a chi-square test because one or more of the cells had a frequency of five or less. For example, when comparing mothers' age by DHS group, there was only one mother aged 40-49 years among the excluded households.

Different statistical tests reflect different facets of validity. It has been argued that performing multiple tests that reflect different facets of validity is superior to performing any single test. (125) The comparison of key socio-demographic characteristics for participating and missing households was therefore followed by a series of three inter-method comparisons aimed at determining the criterion validity

---

[13] The Mann-Whitney U test is a non-parametric analogue to the independent samples t-test that is used when the dependent (ordinal) variable is not normally distributed. Its purpose is to compare the medians for two independent groups.

[14] The Fisher Exact test is used to compare categorical variables when at least one of the categories has an n<5, in lieu of the chi-square test.

of the AWC-IAR and 24HR methods, in comparison to OBS (the gold standard). Terms of measurement properties, for example "validity", "reliability", "reproducibility", "accuracy", "repeatability", etc., are commonly used interchangeably in the literature. (110) In this study, we use the definitions provided by Mokkink et al. (2010) for "criterion validity" ("The degree to which the scores of an ... instrument are an adequate reflection of a ‘‘gold standard") and "reliability" ("The extent to which scores for [participants] who have not changed are the same for repeated measurement under several conditions: for example ... by different persons on the same occasion (interrater)..."), and Altman and Bland (1983) for "agreement".[15]

Inter-method comparison 1: Comparisons of the distributions obtained via OBS v. 24HR or AWC-IAR.

The first inter-method comparison aimed to compare the distributions obtained via OBS versus 24HR, or AWC-IAR using the Wilcoxon[16] signed rank sum for DDS, time allocation and concurrent activities, and McNemar’s test[17] for MDD. Wilcoxon signed rank sum was used instead of a paired t-test because the DDS and time-use data were not normally distributed. The medians of the differences (24HR minus WFR and AWC-IAR minus WFR) were computed, and the distributions of the median differences were also compared.

Comparisons of the distributions obtained via OBS v. 24HR or AWC-IAR were also made at the food group level (i.e., grains, roots and tubers; legumes and nuts; dairy products; flesh foods; eggs; vitamin A-rich fruits and vegetables; and other fruits and vegetables for children; grains, white roots and tubers, and plantains; pulses; nuts and seeds; dairy products; meat, poultry and fish; eggs; dark green leafy vegetables; other vitamin A-rich fruits and vegetables; other vegetables; other fruits for women) and activity level using the Wilcoxon signed rank sum. (See Appendix 16 for the time-use activities used in this study.)

---

[15] As used by Bland and Altman, the definition of "agreement" is that two methods give the same measurement, wherein “the same” means “not significantly different”. (125)

[16] The Wilcoxon signed rank sum is used to compare the medians of two related observations.

[17] McNemar's test is used to compare proportions for two dependent groups.

Inter-method comparison 2: Comparisons of the inter-tool agreements between OBS and 24HR or AWC-IAR using the Bland–Altman limits of agreement method

The Bland-Altman approach is used when assessing agreement[18] between two different methods of measuring some quantity, for example for the purpose of replacing an established method with a new method. (109,126) In this study, we are comparing the AWC-IAR (i.e., the "new" method) to OBS (i.e., the "established" method") of measuring dietary diversity and time use. Historically, comparisons of agreement between a new method and an established one were made using correlation. However, in their seminal 1986 article, "Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement", J. M. Bland and D. G. Altman demonstrated that the use of correlation (or regression analysis) for these purposes is misleading and inappropriate, as evidenced by the following: (1) perfect correlation may be achieved if the measurements of two methods lie along a straight line, whereas perfect agreement is only achieved if the measurements lie along a line of equality; (2) a difference in the scale of the measurement does not affect the correlation, but it affects the agreement; (3) the correlation will be higher when the range of the true quantity of the sample is wide rather than narrow, whereas agreement is not influenced by the range of the true quantity in the sample; (4) the test of significance is irrelevant to the question of agreement; and (5) data which are in poor agreement can produce high correlations. (109,126) Similarly, as a measure of correlation, the Lin's concordance correlation coefficient would not be an appropriate measure of agreement.

Instead, to investigate the relationship between the measurement error and the “true” value, Bland and Altman called for visually inspecting a plot of the difference (method A – method B) against the mean (mean of method A and method B). (115) In a Bland-Altman plot (BAP), the mean of the differences is a measure of systematic bias, whereas the 95% limits of agreement (LOA; mean difference ±2 SD of the differences) are a measure of random error. The shape of the plot indicates how the systematic bias and the amount of random error changes for participants on

[18] As used by Bland and Altman, the definition of "agreement" is that two methods give the same measurement, wherein “the same” means “not significantly different”. (126)

the lower end of the measured outcome versus those at the higher end of the measured outcome. For example, a fan-shaped plot that opens to the right indicates that the random error increases as the measure (e.g., DDS or number of minutes allocated) increases. The Bland-Altman LOA approach was used to assess inter-method agreement for estimating the median DDS, median time allocation per ICATUS activity groups, and the median number of concurrent activities.

Inter-method comparison 3: Comparisons of the inter-tool reliability between OBS and 24HR or AWC-IAR using Cohen's kappa and Cronbach's coefficient alpha

To determine reliability in criterion validity studies, where the dependent variable is ordinal, de Vet et al. (2011) recommends using weighted kappa. (111) Typically, Cohen's kappa is used to assess consistency between multiple *raters* on the same data set. In our study, however, rather than assessing "the extent to which scores for [participants] who have not changed are the same for repeated measurement ... by different *persons* on the same occasion (interrater)...", we assess the reliability between different *methods*, i.e. inter-tool reliability. (110) That is, for this analysis, I used Cohen's kappa to assess the reliability between the gold standard (OBS) and the AWC-IAR or 24HR methods when measuring the same phenomenon. Previous studies have also used Cohen kappa in this way, as a measure of inter-tool reliability. (69,127)

Similarly, Cronbach's coefficient alpha was used as the measure of reliability for the time use data. The intraclass correlation coefficient (ICC) is recommended by de Vet et al (2011) for determining reliability in criterion validation studies where the dependent variable is continuous. (111) ICC quantifies the proportion of the variance of the measurement that is due to the difference between methods, compared to overall variance of the measurement, which in our study was time-use. However, due to the difficulty in interpreting ICC for the purposes of this study, we chose instead to use Cronbach's coefficient alpha as a simple statistical measure of agreement.

The results were interpreted against benchmark scales. Regardless of the outcome measured, Cohen's kappa is conventionally interpreted as follows: <0·00 poor agreement; 0·00–0·20 slight agreement; 0·21–0·40 fair agreement; 0·41–0·60 moderate agreement; 0·61–0·80 substantial agreement; 0·81–1·00 almost perfect agreement. (128,129) Similarly, Cronbach's coefficient alpha is conventionally interpreted as follows: <0.70 unacceptable; >0.70 acceptable; >0.80 moderate; 0.90-0.95 high; >0.95 suspect. (130)

Supplementary validation analyses

To assess whether we could further reduce participant burden by forgoing the IAR, we compared the distributions of DDS and time allocated to different ICATUS activity groups obtained via the criterion method (OBS) versus those obtained when the enumerator coded the images prior to administering the IAR (i.e., the AWC-based enumerator image interpretation) using the Wilcoxon signed rank sum test. Differences in results between OBS and AWC-based enumerator image interpretation or AWC-IAR would be attributable to the additional benefit of engaging participants in the interpretation of the image data.

Also, to assess the possible influence of the observer on the foods/beverages consumed and/or activities performed, DDS and time allocations collected via AWC-IAR were calculated for the non-observation day and compared with those of the corresponding observation day, using the Wilcoxon signed-rank sum (DDS, time allocation, concurrent activities) and McNemar's tests (MDD). We also compared these differences for households that were administered OBS with AWC-IAR before the AWC-IAR without an observer versus households that were administered OBS with AWC-IAR after the AWC-IAR without an observer.

Finally, the frequency of DDS and time allocation differences, and OBS vs. 24HR or AWC-IAR correlations were also calculated.

ETHICAL CONSIDERATIONS

Ethical approval was obtained from the Uganda National Council for Science and Technology (A24ES), the London School of Hygiene & Tropical Medicine Observational/Interventions Research Ethics Committee (project ID: 1420) and the University of Greenwich Faculty of Engineering and Science Ethics Committee (project ID: B0501). The data collection protocols followed the ethical guidelines for AWC research to ensure that privacy of the participants was maintained. (106) See Appendix 17 for ethical approval letters.

Mothers who were randomly selected to participate in the study were notified a day in advance (by Sasakawa community-based facilitators) of the sensitisation meeting in order to give them the chance to discuss participation with their family. Upon notification, they were also encouraged to bring with them to the sensitisation meeting friends or family with whom they might want to discuss participation prior to signing the informed consent form. Only after the community sensitisation, including demonstration of all equipment and devices used in the study (see the Recruitment and Sensitization section above), was written consent or thumb print obtained from mothers who choose to participate in the study.

Each AWC data collection day, participants were reminded that they were free to remove or cover the device at any time. Participants were also given multiple opportunities to delete some or all images.

Chapter 4: Automated wearable cameras for improving recall of diet and time use in Uganda: A feasibility study

ABSTRACT

Background: Traditional recall approaches of data collection for assessing dietary intake and time use are prone to recall bias. Studies in high- and middle-income countries show that automated wearable cameras are a promising method for collecting objective health behaviour data and may improve study participants' recall of foods consumed and daily activities performed. This study aimed to evaluate the feasibility of using automated wearable cameras in rural Eastern Ugandan to collect dietary and time use data.

Methods: Mothers of young children (n=211) wore an automated wearable camera on 2 non-consecutive days while continuing their usual activities. The day after wearing the camera, participants' dietary diversity and time use was assessed using an image-assisted recall. Their experiences of the method were assessed via a questionnaire.

Results: Most study participants reported their experiences with the automated wearable camera and image-assisted recall to be good (36%) or very good (56%) and would participate in a similar study in the future (97%). None of the eight study withdrawals could be definitively attributed to the camera. Fifteen percent of data was lost due to device malfunction, and twelve percent of the images were "uncodable" due to insufficient lighting. Processing and analysing the images were labour-intensive, time-consuming, and prone to human error. Half (53%) of participants had difficulty interpreting the images captured by the camera.

Conclusions: Using an automated wearable camera in rural Eastern Uganda was feasible, although improvements are needed to overcome the challenges common to rural, low-income country contexts and reduce the burdens posed on both participants and researchers. To improve the quality of data obtained, future automated wearable camera-based image assisted recall studies should use a structured data format to reduce image coding time; electronically code the data in the field, as an output of the image review process, to eliminate ex post facto data entry; and, ideally, use computer-assisted personal interviews software to ensure

completion and reduce errors. In-depth formative work in partnership with key local stakeholders (e.g., researchers from low-income countries, representatives from government and/or other institutional review boards, and community representatives and local leaders) is also needed to identify practical approaches to ensuring that the ethical rights of automated wearable camera study participants in low-income countries are adequately protected.

RESEARCH PAPER COVER SHEET

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

<u>SECTION A – Student Details</u>

Student ID Number	1407282	**Title**	Ms.
First Name(s)	Andrea		
Surname/Family Name	Spray		
Thesis Title	Validation of an automated wearable camera-based image-assisted recall method and the 24-hour recall method for assessing maternal and child dietary diversity and women's time use in rural Uganda		
Primary Supervisor	Elaine Ferguson		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

<u>SECTION B – Paper already published</u>

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

<u>SECTION C – Prepared for publication, but not yet published</u>

Where is the work intended to be published?	Nutrition Journal (NutrJ)
Please list the paper's authors in the intended authorship order:	Andrea L. S. Bulungu, Luigi Palla, Joweria Nambooze, Jan Priebe, Lora Forsythe, Pamela Katic, Gwen Varley, Bernice D. Galinda, Nakimuli Sarah, Kate Wellard and Elaine. L.

	Ferguson
Stage of publication	**Undergoing revision**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	A. L. S. B. contributed to the study design and tool development, coordinated data collection and contributed to study management, developed data entry tools, entered data and oversaw data entry research assistants, carried out the feasibility and acceptability analyses, interpreted the results, and wrote and edited the manuscript.

SECTION E

Student Signature	[redacted]
Date	11/11/2022

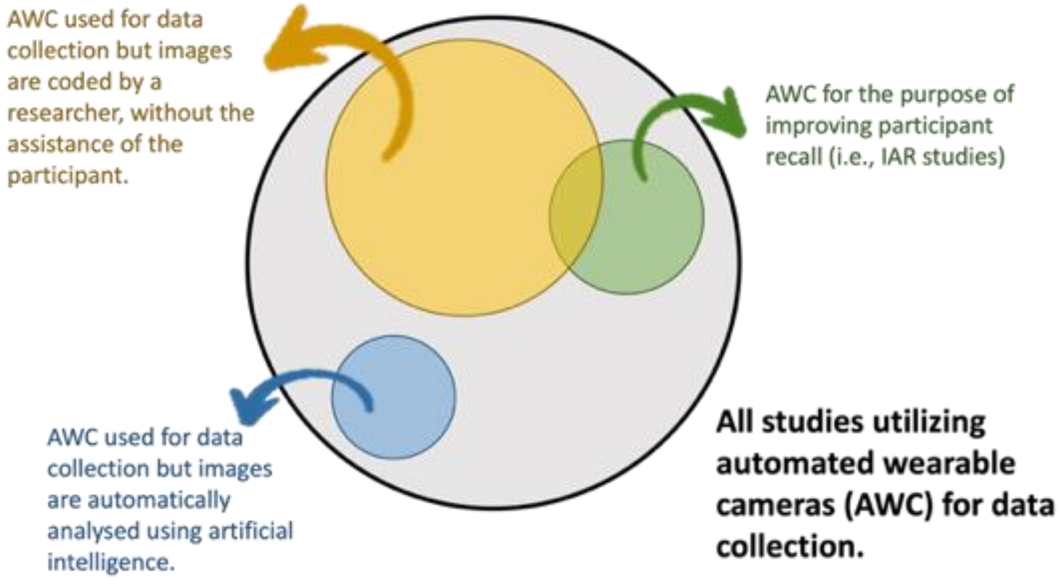
Supervisor Signature	[redacted]
Date	11/11/2022

BACKGROUND

Traditional recall approaches of data collection for assessing dietary intake and time use are prone to recall bias. (13,24,131) Prospective methods, which avoid recall bias, such as self-report diaries are not practicable in rural low-income country contexts due to low literacy, whereas direct observation is labour intensive. An alternative prospective approach is the use of automated wearable cameras. These devices are inexpensive technologies that prospectively and unobtrusively record activities as they are performed. Automated wearable cameras have been used to collect human behaviour data in middle- and high-income countries, but their feasibility in rural, low-income country settings has not yet been determined.

Automated wearable cameras have been evaluated in middle- and high-income countries as a method for improving individuals' recalls of dietary intakes (i.e., food and beverage consumption, eating episodes, and energy intakes), (28,30,32–34,36,67,132,133) the food environment (i.e., food and beverage marketing exposure, purchase, and consumption context), (32,33) and time allocated to daily activities. (47–49) In studies using automated wearable cameras, the captured images have been coded by topical experts, (29,37,39–46,50–52,56–61,64–66,68,105,134) artificial intelligence, (53) or an enumerator with the assistance of the participant via an image-assisted recall. (28,31–36,47–49,55,62,67,135,136) (Figure 1) In an image-assisted recall, photographs which have been taken either automatically via an automated wearable camera or by the participant themselves are used as an memory cue (i.e., recall trigger) to help respondents reconstruct key details from their previous day. (47,137–139) Most image-assisted recall studies provide participants the opportunity to review and delete the images captured by the device privately, before being viewed by the researchers. (106)

Figure 1. Illustration of the relationships between studies utilizing AWCs for data collection



The "feasibility" of automated wearable cameras for collecting health data comprises an array of perceived and objective measures. Perceived measures include the emotional burden on study participants and the people they interact with; ease-of-use of the device; acceptability of image content captured by the wearable camera; and utility of the images captured by the wearable camera in aiding recall. Objective measures include participation refusal, non-compliance, and study withdrawal; device malfunction; observed or reported interactions regarding the camera with members of participants' households and communities; image quality and fit for purpose; time and other resources required for image processing, coding, and analysis; and device cost. Feasibility issues can further be categorized by audience, i.e., from a participant, community, and/or researcher perspective.

Feasibility from a participant perspective

In several studies, some participants found the automated wearable camera cumbersome to wear, especially during physical activity. (28,31,32,45,47,52,54,57,62,63) Studies in which participants were responsible for operating the device (e.g., turning the automated wearable camera on and off at the start and end of the data collection day), commonly reported that participants forgot to wear or charge the device, (45,55,57,59,60,105) or had difficulty pressing the devices' small buttons. (30,51,54,63) Participants have also reported emotional discomfort due to wearing the device, especially in public. (33,34,41,45,63) Heightened awareness of an automated wearable camera may result in a reactive change in the study's behaviour of interest. (28,29,34,41,45,52) In six studies participants reported having modified their behaviour in reaction to being recorded.

Concerns about either wearing the camera or what it might capture may also negatively influence the rates of recruitment and completion. Response rates varied substantially across automated wearable camera-based studies (16% to 89% where reported). (28,31,33,45,52,59) Several of the studies, which explicitly investigated the impact of an automated wearable camera on response rates, attributed recruitment challenges to the device. (33,45,52,59,63,140) Study withdrawal (31,32) and non-compliance (47,52,57) have also been attributed to the use of automated wearable cameras.

For studies using the image-assisted recall method, participants across all age groups reported that viewing the images captured by the automated wearable camera helped them to recall pertinent details of the data collection period. (28,30,32–34,47,67) Participants reported that neither the length of time nor the process of reviewing their automated wearable camera-captured images (i.e., the image-assisted recall) was onerous. (28,47)

Among the 28 studies reporting on automated wearable camera feasibility from a participant perspective, only three were conducted outside of high-income country contexts. (67–69) The evidence from these three studies, which were conducted in middle-income countries, is sparse but consistent with the results already reported. There were no issues related to recruitment or retention, and neither the automated wearable camera nor the image-assisted recall was overly burdensome, however the battery life of the device was insufficient.

Feasibility from a community perspective

Study participants reported removing or covering the automated wearable camera at school, (32,41,47,48) work, (47) home, (52,57) and in public. (62) Three studies reported participants being approached about the automated wearable camera by members of the public, but they were not requested to remove it. (54,55,63)

No studies outside of high-income country contexts have assessed the feasibility of using an automated wearable camera from the community perspective.

Feasibility from a researcher perspective

Lost data due to device inoperability (e.g., insufficient battery life or another malfunction) is among the most commonly reported challenges to the use of automated wearable cameras as a research method. (28–31,33,34,36,39,41,42,46,52,53,57,59,141) Reported data losses due to device inoperability, as a proportion of intended image capture, ranged from 11-50%. (30,31,33,34,38,42,58,59) Most studies report that the images generated by the automated wearable camera are of sufficient quality to enable analysis for the intended purpose. However, several image quality issues are commonly reported across a variety of contexts, including sub-optimal camera angle and positioning, inadequate image capture frequency, and key events that occur off-camera; (28,29,32–34,36,39,42,43,50,57,64,134) dark images caused by low or artificial lighting or obscured lens; (30–32,34,37,41,42,54–57,59–61,64,65) and blurry or scrambled images. (29,37,40,42,57,58,64–66) Furthermore, automated wearable camera images have been reported to be unsuitable for detailed analyses, for specific research areas, such as determining specific items of clothing worn by children far away from the camera, (64) or detecting low intensity activities (e.g., fidgeting or activities performed while sitting down). (50) The proportion of automated wearable camera images reported to be "uncodable" ranged from 1-35%. (37,40–42,44,46,51,52,54,56,57,60,61,64–66)

The results were again sparse but consistent for the two studies conducted in middle-income countries with results related to automated wearable camera feasibility from a researcher perspective. (68,69) In these studies, the cause of data losses was indeterminant and, although the image quality was acceptable, in one study the images captured were unfit for research purpose (determining the quantity of food consumed). (68)

The use of automated wearable cameras for research data collection aims to maximize reporting accuracy while minimizing participant burden. Part of this burden is shifted to the research team, and several studies highlight the heavy time burden required to manually code the automated wearable camera images for analysis, (31,33,39,41,42,45,47,56,61,65,134) and its susceptibility to human error. (45,56) Not all studies quantified the amount of time entailed, but where reported, the estimated time required to code automated wearable camera images range from approximately 1 to 2 hours per participant day. (39–42,44,46,47,51,56,57,59,60,64,65,67) Little information on other costs of automated wearable camera-based research is available. Only Kelly, et al. (2015) reported on the cost of the device (Autographer, £300 each), adding that it was "resource intensive". (47)

Although many studies acknowledged some feasibility limitations, especially for use in large-scale studies, nearly all concluded that automated wearable cameras are a promising method for collecting objective health behaviour data in a free-living setting. Furthermore, despite the challenges described above, studies in high-income countries provide evidence that automated wearable cameras may help to improve study participants' recall of foods consumed, (36) and daily activities performed. (47)

The available evidence for automated wearable camera feasibility, however, almost exclusively derives from studies conducted in high-income and upper-middle income countries. Key characteristics of rural women residents of low-income countries, such as literacy, exposure to technology and social norms, are quite different compared to any of the populations targeted in the automated wearable camera research published thus far. The research environment in rural low-income countries also poses different challenges including, for example, limited access to electricity for lighting the activity space or charging devices, and higher chance of device exposure to dirt or liquids, and lack of enumerators having pertinent skills. Feasibility needs to be explored in low-income countries, especially in rural contexts, given that the environmental conditions (which may affect device operability and image quality), social norms (which may affect acceptability by participants and the public), and familiarity with technological devices (which may affect ease-of-use) are all quite different than in high-income and/or upper-middle-income countries.

This study was therefore undertaken to assess the feasibility of using an automated wearable camera for data collection, in rural Eastern Uganda, on the dietary practices of women and young children and time-use patterns of women. The results can inform future automated wearable camera studies conducted in similar contexts.

Methods

This study is presented per the Strengthen the Reporting of Observational Studies in Epidemiology (STROBE) protocol. (142)

Study design

This study was nested within a cross-sectional study of women with a child aged between 12 and 23 months inclusive (n=211), to examine the impact of a labour-saving technology on women's time for childcare, food preparation and dietary practices. It was conducted between January and February 2018 in Bugiri and Kamuli Districts, Eastern Region, Uganda. It validated the use of three methods of collecting data on dietary practices and women's activities, which were the automated wearable camera-based image-assisted recall, interactive voice response collected via a mobile telephone, and 24-hour recalls, using direct observation as the reference method. Only results related to the automated wearable camera-based image-assisted recall are reported here.

In our study, maternal and child dietary diversity and women's time allocation were assessed via an image-assisted recall using photos captured the previous day with an automated wearable camera. The methods are described in detail elsewhere. (131,143) However, in brief, for each respondent, dietary intake and time allocation data were prospectively collected using photographs automatically taken every 30 seconds by a wearable camera attached to the participant. The next day, using the photos captured by the automated wearable camera during the previous day, an enumerator first independently coded the images for foods / beverages consumed by the mother and child and activities performed by the mother. Then, the enumerator administered an image-assisted recall to the participant. On the day before data collection began, a structured socio-demographic questionnaire was administered, and anthropometric measurements were made. On the final day of data collection, a structured questionnaire was administered to assess participants' perceptions of the automated wearable camera-based image-assisted recall method. Each participant wore the automated wearable camera for two non-consecutive days and completed two image-assisted recalls, which meant enumerators met participants on a total of four days to collect two days of data on dietary practices and women's activities.

Ethical approval was obtained from the Uganda National Council for Science and Technology (UNCST) (A24ES), the London School of Hygiene & Tropical

Medicine Observational Research Ethics Committee (Project ID: 1420), and the University of Greenwich Faculty of Engineering and Science Ethics Committee (Project ID: B0501). Community sensitization was done to ensure the study participants and other community members understood the study objectives and data collection methods. It included a review of key aspects of informed consent, demonstration of the devices used in the study (i.e., automated wearable camera, mobile phone, and GPS tracker), and a detailed description of the methods that would be used in the study, and time was allowed for questions. Following community sensitization, written informed consent (signature or thumb print) was obtained from all respondents who participated in the study.

Participants and sampling

Twelve mother-child dyads were randomly selected from 22 purposefully selected villages in two districts of Eastern Region Uganda, as described elsewhere. (131) Mother-child dyads were excluded if the child was less than 12 months or greater than 23 months of age, was not yet eating solid foods on a regular basis, or was a multiple-birth child; the mother was unable to communicate in Lusoga, Luganda or English; either the mother or child had a severe disability; the mother was not the biological mother of the child; the mother was a co-wife with a selected mother; or either the mother or child was not available for the duration of the study. Participants were given a bar of soap, one kilogram of sugar, a half-litre of cooking oil and a t-shirt on the final day of the study. Also on the final day, they were given a photo of their family taken by a supervisor using a polaroid camera.

Instruments and protocol

The enumerators administered two structured questionnaires to the respondent. The first questionnaire collected information on household socio-demographics and assets, and factors related to women’s empowerment. The second questionnaire, which was administered on the final day of data collection, collected information on household mobile phone access and ownership, and participants' perceptions of their experiences with the automated wearable camera-based image-assisted recall and other data capture methods assessed in this study that are not reported here (i.e., direct observation, 24-hour recall, and mobile phone-based interactive voice response). Specifically, participants were asked to rate the automated wearable camera-based image-assisted recall method using a 4-point Likert scale (very bad,

bad, good, or very good). Participants were also asked to select their favourite and least favourite method among the four data capture methods assessed, and whether they would be willing to participate in an automated wearable camera-based image-assisted recall study again. Although not specifically requested, any comments provided by the participants in answering these questions were translated and transcribed by the enumerators. A brief "innovative methods' questionnaire was also completed at the end of each data collection day to assess participants' experiences wearing the automated wearable camera, including any technical issues or reactions from members of their households or communities. Each participant was also asked to reconfirm her consent to use the images captured by the automated wearable camera. No data on automated wearable camera acceptability among other members of the household or community were collected.

As described elsewhere, (131) a small, lightweight, automated wearable camera (iON SnapCam Lite, dimensions 42 x 42 x 13 mm) was attached to a t-shirt worn by the respondent at approximately 06:00 and removed at approximately 21:00. Participants were instructed to wear the automated wearable camera while continuing their usual activities, and to cover or remove the camera as needed for privacy. A bespoke plastic clip using a safety pin was designed to keep the device firmly attached at the neckline of the t-shirt and well-positioned to minimize interference with clothing. (Figure 2) The wearable camera automatically recorded a picture every 30-seconds, storing all photos (approximately 1,500) on a micro-SD memory card and with the image number (e.g., 4) as the filename (e.g., SNAP0004.JPG). Examples of the photos obtained by device are provided in Supplementary Figure 1.

Figure 2. The AWC affixed via a bespoke clip to the neckline of participants' clothing

The automated wearable camera was turned on at the beginning of the day and turned off at the end of the day by the enumerator. The t-shirt was provided by the study and worn by the participant over her clothing, so that if the participant needed to remove the camera, she would remove the entire t-shirt rather than handling the device. For administrative purposes, at the start of each data collection day, the enumerator took a single picture of the participant, her child and a placard displaying her study ID using the designated function of the device. After attaching the automated wearable camera to the project-provided t-shirt, the enumerator reminded the participant of key points covered during sensitization, i.e., that at any time during the study she could remove or cover the device or request all images to be deleted; that the device was splash proof but could not withstand immersion in water; and to do exactly the activities she would have normally done. Upon picking up the automated wearable camera at the end of the data collection day, the enumerator completed the innovative methods questionnaire. In addition, the first author (ALSB) kept records of inoperable devices, and members of the data collection team monitored issues (e.g., negative rumours about the automated wearable cameras) that may have affected study participation or compliance.

Upon collection of the devices used by the study participants, ALSB saved a copy of the images recorded on each automated wearable camera memory card to a local drive, and assigned two participants' memory cards (i.e., data for two participants) to an enumerator who had not been engaged in direct observation of the participant the previous day. The following day, the enumerator inserted the assigned memory card for the first participant into a tablet (16GB Samsung with a 10” screen, using Simple Gallery software for image display) to review the photos captured by the automated wearable camera. Using the image-assisted recall form, the enumerator annotated the foods she thought were consumed and the activities she thought were undertaken by the participant, and their corresponding image numbers, based on what she could see in the photos. Based on her interpretation, the enumerator demarcated the series of foods and activities for review later that day with the respondent. After completing the annotation for the first assigned participant, the enumerator completed the same steps for the second assigned participant.

Upon meeting with the participant, the enumerator oriented the mother to the photos captured by the automated wearable camera by viewing, on the tablet with the inserted memory card, five pre-selected images: a picture of the mother herself, a picture of her child, a picture of her home, a picture of her garden, and a picture where her own hand is visible while performing a task (e.g., while preparing or cooking food, digging, or using a mobile phone). The enumerator rated the participant's ability to recognize the content of these five photos on a three-point scale: recognized, recognized with help, or failed to recognize. The enumerator then administered the image-assisted recall. During this interview, the enumerator used "verbal probing" (119,120) to elicit from the participant additional relevant information about the activities performed, for example to elaborate on what she was doing, who she was with, where she was going and why, etc. The enumerator revised her original annotations of foods / beverages consumed and activities undertaken by the participant, as needed, based on the participant’s feedback.

The image-assisted recall protocol was adapted from one described by Kelly et al. (2015). (47) The protocol followed ethical guidelines for automated wearable camera research to ensure privacy of the participants was maintained. (106) All protocols were pilot tested and refined prior to the start of the study.

Enumerator training for all devices, protocols, and instruments took place over one week (December 18-22, 2017). The training comprised classroom training, role-play practice, and an assessment with individualized feedback. Training also included two days of field practice.

Data analysis

In this study, feasibility was assessed using administrative data (non-compliance and withdrawal; camera malfunction; image quality; researcher time allocated to data coding and analysis); participants' self-reported ratings of their experiences with the automated wearable camera and other methods used in the study; enumerators' ratings of participants' ability to interpret the images captured by the wearable camera; non-technical (e.g., fear of health or spiritual harm caused by the automated wearable camera) and/or technical (e.g., depleted battery) issues regarding the wearable camera reported by study participants or members of the data collection team; and requests by participant to delete wearable camera captured images.

Participants' self-reported experiences with the automated wearable camera and method ratings were double entered via EpiData. Administrative and demographic data and participant image-assisted recall orientation ratings were entered via Excel. Information about the data processing and analysis of demographic, dietary diversity and time-use data has been previously published. (131,144)

Because the data were not normally distributed, the Mann-Whitney U test and Fisher's Exact test were used to compare method ratings for participating households and households lost to the study. Data were analysed using Stata/SE version 17. P-values less than 0.05 were considered significant for all tests.

RESULTS

Characteristics of the sample

Overall, 211 women were recruited into the study. Among those recruited, twenty-seven participants were eliminated from analysis due to incomplete image-assisted recall data (n=27), including eight participants who voluntarily withdrew from the study (n=8) and two participants who were unavailable for the image-assisted recall due to a funeral (n=1) or medical emergency (n=1). (Supplementary Figure 2) The remaining seventeen (n=17) instances of lost data were due to administrative errors (e.g., inadequate number of tablets or other enumerator-caused image-assisted recall non-complete). No differences in demographic data were found between participants who were excluded from the analyses and those who were analysed. (Supplementary Table 1) The median household size was six members, and approximately one-fifth of participating households lived below $1.25/day. The mean age of participants was 26 years. Most participants were married and had attended primary school. Only half of participants were literate. Slightly more than half of the participants identified as Christian.

Feasibility from a participant perspective

Most participants rated their experience of wearing the automated wearable camera and reviewing the photos the following day (image-assisted recall) as either good (36%) or very good (56%). (Supplementary Table 2) For over a quarter of participants (26%), the automated wearable camera-based image-assisted recall was their favourite of the four research methods assessed, which was significantly higher than the proportion who preferred the 24-hour recall method (4%). (Supplementary Table 3a - Supplementary Table 3d) These participants reported that they enjoyed looking at the photographs and that the photographs helped them remember key details.

For over a quarter (29%) of study participants, the automated wearable camera-based image-assisted recall was their least favourite of the four research methods assessed. For these participants, invasion of privacy, fear of the device and fear of others' reaction to the device (emotional burden) were contributing factors. This result was not significantly different than the proportions who rated direct observation (24%), 24-hour recall (32%), or the mobile phone-based interactive voice response (15%) as their least favourite method. (Supplementary Table 4a - Supplementary Table 4d)

Nearly all (96%) participants reported that they would be willing to wear an automated wearable camera to record their food intake and daily activities in a future study. Furthermore, none of the eight study withdrawals can be definitively attributed the automated wearable camera. In three cases, the participant's husband declined after she had initially consented. One participant withdrew after expressing frustration with one of the other data capture methods (i.e., mobile phone-based interactive voice response) assessed in this study ("I am tired, tired, tired of your things"). Two reported being called away to attend a burial and no rationale for withdrawal was provided by the remaining dropouts (n=2). Most (n=6) of the withdrawals occurred in one of the two study districts (Kamuli) which resulted in a higher withdrawal rate in this district than in Bugiri (i.e., 9% in Kamuli vs 1% in Bugiri).

Across all participants, including those who withdrew, none requested that their image data be entirely deleted either at the end of the data collection day or after viewing their photos during the image-assisted recall. Seven participants requested that a few specific photos be deleted. In two cases, the participants were aware that the camera had likely captured private activities (bathing children and using the latrine) and requested pictures to be deleted at the time of device pick-up. In three cases, the participant vaguely indicated they wanted "a few" images deleted but did not specify which images at the time of data collection or during the image-assisted recall. In the remaining two instances, participants requested specific images to be deleted after seeing them in the image-assisted recall.

Feasibility from a researcher perspective

An additional 15% of data (n=27) had an insufficient number of images captured due to inoperability of the automated wearable camera. The causes and/or nature of these automated wearable camera malfunctions were usually unclear. In only about a half (n=16) of these cases were operability issues reported at the time of data collection (usually identified by either the failure of the device to display a blue light indicating a photo had been taken or to beep in response to button pressing) by either the participant or a member of the research team. In five instances the enumerator inadvertently recorded a video, which may have depleted the device battery or storage capacity. There were also instances (n=15) where an operability issue was reported but, if a malfunction occurred, it did not substantially reduce the number of images captured. The number and proportion of inoperable automated wearable cameras increased over the course of the study. (Supplementary Table 5) Only 6% of the study population in Bugiri was affected by inoperable devices versus 26% in Kamuli. Photos taken by the automated wearable camera before dawn, in the evening, or indoors were often too dark to determine foods or activities. Overall, about twelve percent of images captured by the automated wearable camera were too dark to interpret.

Approximately half (53%) of participants had difficulty interpreting at least one of the five pre-selected "orientation" images (i.e., a picture of the mother herself, a picture of her child, a picture of her home, a picture of her garden, and a picture where her own hand is visible while performing a task). Ease-of-recognition of two types of orientation photos improved between the first and second image-assisted recall - the photo of the participant's garden (from 93% to 97%) and the photo of her own hand while performing a task (from 89% to 96%). (Supplemental Table 6)

Discussion

Principal Results

This is the first study to investigate the feasibility of an automated wearable camera-based image-assisted recall for collecting maternal and child dietary diversity or women's time-use data in a low-income country context. We assessed the feasibility of the automated wearable camera-based image-assisted recall method with mothers of young children in rural Eastern Region Uganda using administrative data and participant-reported perceptions. Results showed the collection of food / beverage intake and women's time-use data, using an automated wearable camera, was feasible, although data loss was high.

There were no reports of physical discomfort due to the automated wearable camera. Unlike previous studies that have hung the automated wearable camera on a lanyard, which may swing and get in the way of daily activities, in this study a bespoke clip was used to securely fasten the automated wearable camera to a t-shirt that was large enough to fit over the participant's clothing. Camera malfunction was minimized because participants in this study were not responsible for operating the device. Accordingly, there also were no instances of lost data due to participants forgetting to wear the automated wearable camera or failing to recharge it, nor were any usability issues reported. There were, however, eight instances of inadvertent video recording, possibly indicating a usability issue with the device when operated by the trained enumerator. These cases can lead to lost data because video recording uses more battery and storage space.

Based on the high participant retention rate and the participants' end-of-study method ratings, the acceptability of the automated wearable camera among participants of this study was high. Although a fifth of participants deemed it their least favourite method among those assessed in this study, 92% still rated the automated wearable camera-based image-assisted recall method as good or very good and 97% were willing to participate in an automated wearable camera-based image-assisted recall study in the future.

Although eight participants withdrew from the study, none could be definitively attributed to the automated wearable camera. Three of the eight withdrawals occurred because the participant's spouse refused to participate (after the participant had consented), which underscores the importance of careful sensitization of not only potential participants but also their families and their communities prior to recruitment. Careful monitoring (e.g., for rumours) during the period of data collection is also critical to address any concerns before they become a more widespread problem. This also underscores the importance of having a field team that speaks the local language and is familiar with the local culture, as well as strong engagement with and support of community leaders. Notably, a higher proportion of the participants who withdrew were from Kamuli than Bugiri, which may indicate that social acceptability can vary across small-scale geographies. In small close-knit communities such as the ones where this study was conducted, a single negative event (e.g., illness or death of a community member or family member) or rumours (e.g., that the automated wearable camera causes spiritual harm) can influence participation. Because the automated wearable camera-based image-assisted recall method requires two days of data collection - one day to wear the camera and the next day to review the images captured by the automated wearable camera - it is vulnerable to unanticipated absences. Future research to explore the influence of social networks on automated wearable camera study participation may be warranted.

Participants in this study were offered multiple opportunities to delete all or some of the images captured by the automated wearable camera. None of the participants requested that their data be entirely deleted and only seven participants requested that a few specific photos be deleted. This low level of deletion indicates that participants found the content of the automated wearable camera images acceptable, or that they did not feel comfortable asking the enumerators for the photos to be deleted. Among the seven studies reporting on the acceptability of the content captured by the automated wearable camera (all in high-income countries), six reported participants opting to delete images. (46,52,56,57,62,63) Drawing comparisons between this study and previous studies is difficult because, for most image-assisted recall studies, participants are not obligated to report if or how many images they deleted. The ethical framework outlined by Kelly et al. for the use of automated wearable cameras in health behaviour research (106) recommends that participants are provided time to review their images in private before being viewed by the research team, and to delete any images desired. This approach is impractical in a low-literacy population with limited exposure to digital technologies. This constraint and other recommendations of the Kelly framework warrant review by key local stakeholders, such as researchers from low-income countries, representatives from government and/or other institutional review boards, and community representatives and local leaders, to identify practical approaches to ensuring that the ethical rights of automated wearable camera study participants in low-income countries are adequately protected.

This study was not designed to assess acceptability from the community perspective. In the end-of-day innovative methods questionnaire, however, there were no reported requests from the public to remove the automated wearable camera, although one participant reportedly removed the device after being "threatened by others that she is being recruited for the Illuminati" (supervisor's field notes). Future studies need to investigate acceptability from the community perspective given the substantial socio-cultural differences between high-income country contexts and rural low-income country contexts, and across different low-income countries.

Fifteen percent of data was lost due to automated wearable camera malfunction (n=27), which is similar to reported data losses due to device malfunction in other studies.(30,31,33,34,36,42,58,59) The relatively low rate of device malfunction may be because they were operated by the enumerators rather than the participants, and could be quickly replaced when an enumerator was present on the observation day. Indeed, the number instances of lost data due to inoperability was higher on days when the observer was not present at the home (n=24 and n=32 when the observer was and was not present, respectively). (Supplementary Table 7) Nevertheless, any amount of lost data is a waste of study and participant resources and may decrease the power of the study to detect the outcomes intended. Automated wearable camera operational issues increased over the course of the study, which suggests that wear and tear on the devices or the SD cards rather than inadequate battery charging were at fault. Back-up devices may need to be procured so that malfunctioning automated wearable cameras in the field can be replaced. However, it was not always evident when a device was inoperable until the end of the data collection day when the images were downloaded. Future studies will need to over-recruit participants in anticipation of higher-than-usual attenuation compared with the 24-hour recall or direct observation methods.

Overall, the quality of images captured by the automated wearable camera was sufficient for the purposes of assessing dietary diversity and time use, however, twelve percent of the images were "uncodable" due to insufficient lighting (i.e., too dark to determine foods or activities). This result is within the range of uncodable images reported in previous studies. (37,40–42,44,46,51,52,54,56,57,60,61,64–66) Due to long periods of indecipherable activities at the beginning and end of each day, however, the time use validation study analysis was reduced from 15 to 12 hours. Constraints on the period of data collection limits the comparability of results obtained via automated wearable camera-based image-assisted recall versus other methods, for example, because of missing behaviours of interest that only occur in low-lit contexts, such as storytelling with children before bed or consumption of more nutritious dishes in the evening than during the day. Future studies are advised to explore unobtrusive options for improving ambient light when the participant is indoors or in the morning or evening when sunlight is low.

In addition, automated wearable camera photos captured during vigorous activity, for example while sweeping or riding a bicycle, were sometimes blurry, although for the purposes of this study interpretation was not hindered when viewed in context of the surrounding images. The image quality was sometimes inadequate to determine slight colour variations that are important for dietary assessment, such as the difference between white and yellow sweet potatoes, which would be important in studies where the coding was performed by the researcher without the assistance of the participant. Food consumption and preparation steps may have also been missed between 30-second photo increments. Fruit, in particular, is often picked and consumed quickly, appearing in just 1 or 2 frames (out of approximately 1,500), or not at all. These issues may be addressed with an automated wearable camera with a higher frequency capture rate, although Arab & Winter (2010) reported that foods are still missed even when using an automated wearable camera with 10-second frequency of data capture, (29) and a higher image capture frequency would also increase the number of images that need to be coded, posing an additional burden on the research team.

The automated wearable camera used in this study (iON SnapCam Lite) named the captured images by image number rather than timestamp. Although the timestamp was embedded in the file's metadata, it was not possible to see at-a-glance the time the photograph was taken to determine, for example, the timeslot in which the activity occurred or to easily calculate elapsed time. Furthermore, the automated wearable camera tended to "lose" time over the course of the study. (Each day the automated wearable camera's clock became increasingly out-of-sync with real time.) Therefore, even the timestamp embedded in the image's metadata was not reliable, although it could be used to determine activity duration.

The positioning of the automated wearable camera around the neckline of the mother may also have resulted in missed data related to food/beverage consumption and activities, especially as it pertains to childcare and child feeding. For example, there was no visual record of foods consumed by the child under the care of someone else, potentially nutrient-rich ingredients added to a recipe prepared by another member of the household, passive caregiving when the mother was not facing her child and socializing while simultaneously engaged in other activities such as washing clothes. Future studies should carefully consider the placement of the automated wearable camera given the study objectives.

Finally, some participants had difficulty interpreting first-person perspective photographs captured by the wearable camera. This is notable because the image-assisted recall method is designed to "trigger" the participant's recall of activities, however the images provide no assistance to (or worse, may misdirect) a participant's recall of events if they cannot interpret what they see in the photo. Although participants easily identified the photos of their children and themselves, a substantial proportion struggled to interpret the photos of themselves engaged in a task. This issue has never before been reported, and it's unclear what, if any, predictive value a participant's skill at identifying a single still "action" image may have for their overall image-assisted recall performance. To reduce the burden of the method for participants and researchers in future studies, it would be useful to determine the minimum number of photos and the type of photo (e.g., a well-chosen single image) that is effective for improving recall.

A thorough analysis of the validity of automated wearable cameras and the 24-hour recall methods were reported separately. (3,60)

Strengths

Scalability is an important decision when choosing research methods. This feasibility study was conducted in communities with free-living participants, which reflects the conditions in which research is typically carried out in rural low-income countries. Compared to previous automated wearable camera feasibility studies, which are commonly conducted with populations less than a hundred, this study was conducted with a relatively large number of participants.

This study also reports two innovative and beneficial research design choices. First, participants in this study were not responsible for operating the devices, which, although increasing the burden on the research team, likely reduced lost data. Second, an orientation to the automated wearable camera-captured images was added at the start of the image-assisted recall to help participants with little previous experience viewing first-person photography to use the images to trigger their memories. Future studies should explore whether these practices are effective in reducing reporting error.

Limitations

This feasibility study was nested within a rigorous validation study, which necessitated the concurrent administration of multiple methods (e.g., 15-hour direct observation, 24-hour recalls and interactive voice response). To reduce the burden imposed on participants, the methods used to assess their perceptions of the AWC and IAR were limited to a brief semi-structured questionnaire at the end of each AWC data collection day and a simple survey of closed-ended questions at the end of the 5-day data collection period. Future studies should incorporate more robust qualitative methods to assess "acceptability" of the AWC in rural LIC contexts. For example, that nearly all participants rated the automated wearable camera-based image-assisted recall method as good or very good and were willing to participate in an automated wearable camera-based image-assisted recall study in the future was a surprising result. In-depth qualitative investigation is needed to unpack this finding.

Also due to the concurrent collection of data via multiple methods, the interaction between the study team and participants was more intense in this study than would occur in typical studies using the automated wearable camera. Consequently, lost data was possibly lower in this study than it would ordinarily be because an observer in the home would have more readily identified a malfunctioning automated wearable camera and alerted the research team for quick replacement. Acceptability-related study withdrawals may have been lower in this study because the observers also helped to monitor negative rumours so the research team could address them early. However, acceptability may have been adversely affected because of high participant burden due to multiple simultaneous methods.

This study was conducted in a single population and so the results cannot be generalized to other populations, even to other rural populations in Uganda. Participants in this study spent most of their time at home. The feasibility of an automated wearable camera may be lower for participants who work away from home. The feasibility of the image-assisted recall, which took approximately 1-2-hours to administer, may also be lower in periods when workloads are heavy, and participants are experiencing time constraints e.g., during the planting or harvest seasons.

Although the lack of significant differences in the household characteristics among those included and excluded from the study indicates that neither selection bias nor self-selection bias diminished the internal validity of the results among those recruited, the possible effect of the automated wearable camera on recruitment was not investigated. Future studies should try to monitor the reasons why people refuse to participate. For example, to better understand the factors that had influenced recruitment rates, non-participants could be asked to provide the reason, as done by Cowburn et al. (2015). (33) Research to assess the perceptions of study participants' family and community members regarding the automated wearable camera is also needed, as is research to estimate the cost-effectiveness of the automated wearable camera-based image-assisted recall method versus traditional recall methods, accounting for all equipment costs (including but not limited to the automated wearable cameras) and researcher time required.

Conclusions

In conclusion, this study showed that using an automated wearable camera in rural Eastern Uganda was feasible. The results inform future studies about investments to improve feasibility in these contexts, including the need for higher quality devices, more automated data management processes, and more in-depth sensitization of study participants, their families, and communities.

The most critical issues were those that resulted in substantial lost and/or unusable data, e.g., automated wearable camera malfunction, poor image quality, and poor device usability. Although these challenges may be alleviated by an investment in higher quality cameras, they are commonly reported in other automated wearable camera studies using top-of-the-line devices explicitly designed for the purpose of behavioural research data collection.

Coding the automated wearable camera images for analysis was also resource intensive, a factor that must be considered in weighing the trade-offs between different data collection methods. This challenge, too, is commonly reported in other automated wearable camera studies. Rough estimates calculated ex post facto for this study, indicate that automated wearable camera image coding took approximately five researcher hours per participant day, substantially more than previously reported. Although not insurmountable, innovations engaging stakeholders in a wide array of fields (e.g., computer science, engineering, and social science) are needed to improve hardware, software, and data analysis methods for the automated wearable camera-based image-assisted recall method to be scalable, regardless of country context.

More unique to rural, low-income country settings, and therefore warranting special consideration for future studies in these contexts, are the lack of exposure to first-person photos, insufficient experience operating a computer (e.g., to independently review automated wearable camera-captured images), and different socio-cultural norms compared to the contexts in which current automated wearable camera protocols have evolved. In this study, separate tools were used by the participant to review their photos (a tablet computer) and by the enumerator collect the recall data (a paper-based, unstructured instrument). The image-assisted recall data was transcribed verbatim in the field and coded into the designated categories ex post facto. To improve the quality of data obtained via automated wearable cameras, future image-assisted recall studies should use a structured data format to reduce automated wearable camera image coding time; electronically code the data in the field, as an output of the image review process, to eliminate ex post facto data entry; and, ideally, use computer-assisted personal interviews (CAPI) software to ensure completion and reduce errors. In high-income countries, automated wearable camera-based image-assisted recall researchers have used the SenseCam browser, (145) specifically designed for the purpose of administering image-assisted recalls, to scroll through the images while simultaneously coding them in a controlled environment. Similar applications are needed for rural low-income country contexts where the image-assisted recall is administered at the participant's home (likely without access to electricity) and participants lack the skills to operate a computer. For these challenges, in-depth formative work is needed to specifically design automated wearable camera-based image-assisted recall methods that work for these contexts.

There is no comparable method for capturing such rich and diverse data simultaneously and prospectively, on dimensions the participant may not think to report, or for enabling analysis of new questions that emerge after the data has been collected. Further work to design automated wearable camera-based image-assisted recall protocols is needed, however, to overcome the challenges common to rural, low-income country contexts and reduce the burdens posed on both participants and researchers and data losses.

REFERENCES

1. Black RE, Victora CG, Walker SP, Bhutta ZA, Christian P, De Onis M, et al. Maternal and child undernutrition and overweight in low-income and middle-income countries. Lancet [Internet]. 2013;382(9890):427–51. Available from: http://linkinghub.elsevier.com/retrieve/pii/S014067361360937X

2. Stanaway JD, Afshin A, Gakidou E, Lim SS, Abate D, Abate KH, et al. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Stu. Lancet [Internet]. 2018 Nov;392(10159):1923–94. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673618322256

3. Afshin A, Sur PJ, Fay KA, Cornaby L, Ferrara G, Salama JS, et al. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2019 May;393(10184):1958–72.

4. Horton S, Steckel RH. Malnutrition: Global economic losses attributable to malnutrition 1900-2000 and projections to 2050. Copenhagen: Copenhagen Consensus on Human Challenges; 2011. (Copenhagen Consensus on Human Challenges Assessment Paper).

5. FAO. The State of Food Insecurity in the World 2013: The multiple dimensions of food security. Rome: Food and Agriculture Organization of the United Nations; 2013. 114 p.

6. United Nations Children's Fund (UNICEF), World Health Organization, International Bank for Reconstruction and Development/The World Bank. Levels and trends in child malnutrition: Key findings of the 2021 edition of the joint child malnutrition estimates. Geneva: World Health Organization; 2021.

7. UNICEF. Strategy for improved nutrition of infant and women in developing countries. New York: UNICEF; 1990.

8. Black RE, Allen LH, Bhutta ZA, Caulfield LE, de Onis M, Ezzati M, et al. Maternal and child undernutrition: global and regional exposures and health consequences. Lancet [Internet]. 2008;371(9608):243–60. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0140673607616900

9. World Health Organization (WHO), United Nations Children's Fund (UNICEF). Indicators for assessing infant and young child feeding practices: definitions and measurement methods. Geneva: World Health Organization and the United Nations Children's Fund (UNICEF); 2021.

10. Engle PL, Menon P, Haddad L. Care and Nutrition: Concepts and Measurement. Washington, D.C.; 1997.

11. International Labour Office. Women at Work: Trends 2016. Geneva; 2016.

12. Kes A, Swaminathan H. Gender and Time Poverty in Sub-Saharan Africa. In: Blackden CM, Wodon Q, editors. Gender, Time Use, and Poverty in Sub-Saharan Africa. Washington, D.C.: The International Bank for Reconstruction and Development/The World Bank; 2006. p. 13–38.

13. Ricci J, Jerome N, Megally N, Galal O, Harrison G, Kirksey A. Assessing the Validity of Informant Recall: Results of a Time Use Pilot Study in Peri-Urban Egypt. Hum Organ [Internet]. 1995;54(3):304–8. Available from: http://sfaajournals.net/doi/10.17730/humo.54.3.a536540628381010

14. Robinson JP. The time-diary method: Structure and uses. In: Pentland WE, Harvey AS, Lawton MP, McColl MA, editors. Time Use in the Social Sciences. New York: Kluwer Academic Publishers; 1999. p. 47–89.

15. Gross DR. Time allocation: A tool for the study of cultural behavior. Annu Rev Anthropol. 1984;13:519–58.

16. Bério A-J. The analysis of time allocation and activity patterns in nutrition and rural development planning. Food Nutr Bull [Internet]. 1984 Mar 2;6(1):1–16. Available from: http://journals.sagepub.com/doi/10.1177/156482658400600113

17. Marini MM, Shelton BA. Measuring Household Work: Recent Experience

in the United States. Soc Sci Res [Internet]. 1993 Dec;22(4):361–82. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0049089X83710185

18. Kalton G. Sample design issues in time diary studies. In: Juster FT, Stafford FP, editors. Time, Goods, and Well-Being. Ann Arbor, Michigan: Institute for Social Reearch, The University of Michigan; 1985. p. 93–112.

19. Robinson JP. The Validity and Reliability of Diaries versus Alternative Time Use Measures. In: Juster FT, Stafford FP, editors. Time, Goods, and Well-Being. Ann Arbor, Michigan: Institute for Social Reearch, The University of Michigan; 1985. p. 33–62.

20. Wagenbuur HTM. An analysis of the time-budget and of the production and productivity of Lime Farmers In Southern Ghana. The Hague: Institute of Social Studies; 1972. (I.S.S. Ocassional Papers).

21. Gardner F. Methodological Issues in the Direct Observation of Parent–Child Interaction: Do Observational Findings Reflect the Natural Behavior of Participants? Clin Child Fam Psychol Rev. 2000;3:185–198.

22. Repp AC, Nieminen GS, Olinger E, Brusca R. Direct Observation: Factors Affecting the Accuracy of Observers. Except Child [Internet]. 1988 Sep 24;55(1):29–36. Available from: http://journals.sagepub.com/doi/10.1177/001440298805500103

23. Richer J. Direct Observation: Impediments and Approaches. Hum Ethol Bull [Internet]. 2017 Dec 28;32(4):6–14. Available from: http://ishe.org/human-ethology-bulletin/2017-2/heb-324-special-issue-why-direct-observation/direct-observation-impediments-and-approaches/

24. Bernard HR, Killworth P, Kronenfeld D, Sailer L. The Problem of Informant Accuracy: The Validity of Retrospective Data. Annu Rev Anthropol [Internet]. 1984;13(1):495–517. Available from: http://www.annualreviews.org/doi/10.1146/annurev.an.13.100184.002431

25. Beaton GH, Milner J, McGuire V, Feather TE, Little JA. Source of variance in

24-hour dietary recall data: implications for nutrition study design and interpretation. Carbohydrate sources, vitamins, and minerals. Am J Clin Nutr [Internet]. 1983 Jun 1;37(6):986–95. Available from: https://academic.oup.com/ajcn/article/37/6/986/4690872

26. Poslusna K, Ruprich J, de Vries JHM, Jakubikova M, van't Veer P, Vries JHM De, et al. Misreporting of energy and micronutrient intake estimated by food records and 24 hour recalls , control and adjustment methods in practice. Br J Nutr [Internet]. 2018;101(2009):S73. Available from: http://www.journals.cambridge.org/abstract_S0007114509990602

27. Gibson RS, Charrondière UR, Bell W. Measurement Errors in Dietary Assessment Using Self-Reported 24-Hour Recalls in Low-Income Countries and Strategies for Their Prevention. Adv Nutr An Int Rev J [Internet]. 2017 Nov 15;8(6):980–91. Available from: http://advances.nutrition.org/lookup/doi/10.3945/an.117.016980

28. Arab L, Estrin D, Kim DH, Burke J, Goldman J. Feasibility testing of an automated image-capture method to aid dietary recall. Eur J Clin Nutr [Internet]. 2011;65(10):1156–62. Available from: http://www.nature.com/articles/ejcn201175

29. Arab L, Winter A. Automated Camera-Phone Experience with the Frequency of Imaging Necessary to Capture Diet. J Am Diet Assoc [Internet]. 2010 Aug;110(8):1238–41. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0002822310005304

30. O'Loughlin G, Cullen SJ, Mcgoldrick A, Connor SO, Blain R, Malley SO, et al. Using a Wearable Camera to Increase the Accuracy of Dietary Analysis. AMEPRE [Internet]. 2013 Mar;44(3):297–301. Available from: https://linkinghub.elsevier.com/retrieve/pii/S074937971200863X

31. Beltran A, Dadabhoy H, Ryan C, Dholakia R, Jia W, Baranowski J, et al. Dietary Assessment with a Wearable Camera among Children: Feasibility and Intercoder Reliability. J Acad Nutr Diet [Internet]. 2018 Nov;118(11):2144–53. Available from: https://doi.org/10.1016/j.jand.2018.05.013

32. Kamar M, Evans C, Hugh-Jones S. Factors Influencing British Adolescents' Intake of Whole Grains: A Pilot Feasibility Study Using SenseCam Assisted Interviews. Nutrients [Internet]. 2019 Nov 1;11(11):2620. Available from: https://www.mdpi.com/2072-6643/11/11/2620

33. Cowburn G, Matthews A, Doherty A, Hamilton A, Kelly P, Williams J, et al. Exploring the opportunities for food and drink purchasing and consumption by teenagers during their journeys between home and school: a feasibility study using a novel method. Public Health Nutr [Internet]. 2015 Jan 15;19(1):93–103. Available from: https://www.cambridge.org/core/product/identifier/S1368980015000889/type/journal_article

34. Gemming L, Doherty A, Kelly P, Utter J, Ni Mhurchu C. Feasibility of a SenseCam-assisted 24-h recall to reduce under-reporting of energy intake. Eur J Clin Nutr [Internet]. 2013;67(10):1095–9. Available from: http://www.nature.com/doifinder/10.1038/ejcn.2013.156

35. Beltran A, Dadabhoy H, Chen TA, Lin C, Jia W, Baranowski J, et al. Adapting the eButton to the Abilities of Children for Diet Assessment. In: Proceedings of Measuring Behavior 2016: 10th International Conference on Methods and Techniques in Behavioral Research International Conference on Methods and Techniques in Behavioral Research. Dublin, Ireland: NIH Public Access; 2016.

36. Gemming L, Rush E, Maddison R, Doherty A, Gant N, Utter J, et al. Wearable cameras can reduce dietary under-reporting: doubly labelled water validation of a camera-assisted 24 h recall. Br J Nutr [Internet]. 2015 Jan 28;113(2):284–91. Available from: https://www.cambridge.org/core/product/identifier/S0007114514003602/type/journal_article

37. Smith M, Stanley J, Signal L, Barr M, Chambers T, Balina A, et al. Children's healthy and unhealthy beverage availability, purchase and consumption: A wearable camera study. Appetite [Internet]. 2019;133(July 2018):240–51. Available from: https://doi.org/10.1016/j.appet.2018.11.004

38. Pettitt C, Liu J, Kwasnicki RM, Yang G-Z, Preston T, Frost G. A pilot study to determine whether using a lightweight, wearable micro-camera improves dietary assessment accuracy and offers information on macronutrients and eating rate. Br J Nutr [Internet]. 2016;115(01):160–7. Available from: http://www.journals.cambridge.org/abstract_S0007114515004262

39. Raber M, Patterson M, Jia W, Sun M, Baranowski T. Utility of eButton images for identifying food preparation behaviors and meal-related tasks in adolescents. Nutr J [Internet]. 2018 Dec 24;17(1):32. Available from: https://nutritionj.biomedcentral.com/articles/10.1186/s12937-018-0341-2

40. Raber M, Baranowski T, Crawford K, Sharma S V., Schick V, Markham C, et al. The Healthy Cooking Index: Nutrition Optimizing Home Food Preparation Practices across Multiple Data Collection Methods. J Acad Nutr Diet [Internet]. 2020 Jul;120(7):1119–32. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2212267220300629

41. Barr M, Signal L, Jenkin G, Smith M. Capturing exposures: Using automated cameras to document environmental determinants of obesity. Health Promot Int [Internet]. 2014;30(1):56–63. Available from: https://academic.oup.com/heapro/article-lookup/doi/10.1093/heapro/dau089

42. Gemming L, Doherty A, Utter J, Shields E, Ni Mhurchu C. The use of a wearable camera to capture and categorise the environmental and social context of self-identified eating episodes. Appetite [Internet]. 2015;92:118–25. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0195666315002469

43. McKerchar C, Smith M, Gage R, Williman J, Abel G, Lacey C, et al. Kids in a Candy Store: An Objective Analysis of Children's Interactions with Food in Convenience Stores. Nutrients [Internet]. 2020 Jul 18;12(7):2143. Available from: https://www.mdpi.com/2072-6643/12/7/2143

44. Mckerchar C, Smith M, Stanley J, Barr M, Chambers T, Abel G, et al. Food

store environment examination – FoodSee: A new method to study the food store environment using wearable cameras. Glob Health Promot. 2019;0(0):1–9.

45. Schrempft S, van Jaarsveld CH, Fisher A. Exploring the Potential of a Wearable Camera to Examine the Early Obesogenic Home Environment: Comparison of SenseCam Images to the Home Environment Interview. J Med Internet Res [Internet]. 2017 Oct 12;19(10):e332. Available from: http://www.jmir.org/2017/10/e332/

46. Signal LN, Smith MB, Barr M, Stanley J, Chambers TJ, Zhou J, et al. Kids'Cam: An Objective Methodology to Study the World in Which Children Live. Am J Prev Med [Internet]. 2017 Sep;53(3):e89–95. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0749379717301630

47. Kelly P, Thomas E, Doherty A, Harms T, Burke Ó, Gershuny J, et al. Developing a Method to Test the Validity of 24 Hour Time Use Diaries Using Wearable Cameras: A Feasibility Pilot. Harezlak J, editor. PLoS One. 2015 Dec 3;10(12):e0142198.

48. Gershuny J, Harms T, Doherty A, Thomas E, Milton K, Kelly P, et al. Testing Self-Report Time-Use Diaries against Objective Instruments in Real Time. Sociol Methodol [Internet]. 2020 Aug 30;50(1):318–49. Available from: http://journals.sagepub.com/doi/10.1177/0081175019884591

49. Harms T, Gershuny J, Doherty A, Thomas E, Milton K, Foster C. A validation study of the Eurostat harmonised European time use study (HETUS) diary using wearable technology. BMC Public Health [Internet]. 2019 Jun 3;19(S2):455. Available from: https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-019-6761-x

50. Davies A, Allman-Farinelli M, Owen K, Signal L, Hosking C, Wang L, et al. Feasibility Study Comparing Physical Activity Classifications from Accelerometers with Wearable Camera Data. Int J Environ Res Public Health [Internet]. 2020 Dec 13;17(24):9323. Available from:

https://www.mdpi.com/1660-4601/17/24/9323

51. Doherty AR, Kelly P, Kerr J, Marshall S, Oliver M, Badland H, et al. Using wearable cameras to categorise type and context of accelerometer-identified episodes of physical activity. Int J Behav Nutr Phys Act [Internet]. 2013;10(1):22. Available from: http://ijbnpa.biomedcentral.com/articles/10.1186/1479-5868-10-22

52. Everson B, Mackintosh KA, McNarry MA, Todd C, Stratton G. Can Wearable Cameras be Used to Validate School-Aged Children's Lifestyle Behaviours? Children [Internet]. 2019 Feb 1;6(2):20. Available from: http://www.mdpi.com/2227-9067/6/2/20

53. Gurrin C, Qiu Z, Hughes M, Caprani N, Doherty AR, Hodges SE, et al. The smartphone as a platform for wearable cameras in health research. Am J Prev Med. 2013;44(3):308–13.

54. Harvey JA, Skelton DA, Chastin SFM. Acceptability of novel lifelogging technology to determine context of sedentary behaviour in older adults. AIMS Public Heal [Internet]. 2016;3(1):158–71. Available from: http://www.aimspress.com/article/10.3934/publichealth.2016.1.158

55. Kelly P, Doherty A, Berry E, Hodges S, Batterham AM, Foster C. Can we use digital life-log images to investigate active and sedentary travel behaviour? Results from a pilot study. Int J Behav Nutr Phys Act [Internet]. 2011;8(1):44. Available from: http://ijbnpa.biomedcentral.com/articles/10.1186/1479-5868-8-44

56. Kerr J, Marshall SJ, Godbole S, Chen J, Legge A, Doherty AR, et al. Using the SenseCam to Improve Classifications of Sedentary Behavior in Free-Living Settings. Am J Prev Med [Internet]. 2013 Mar;44(3):290–6. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0749379712008495

57. Smith C, Galland BC, de Bruin WE, Taylor RW. Feasibility of Automated Cameras to Measure Screen Use in Adolescents. Am J Prev Med [Internet]. 2019 Sep;57(3):417–24. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0749379719302065

58. Chambers T, Pearson AL, Kawachi I, Rzotkiewicz Z, Stanley J, Smith M, et al. Kids in space: Measuring children's residential neighborhoods and other destinations using activity space GPS and wearable camera data. Soc Sci Med [Internet]. 2017;193:41–50. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0277953617305907

59. Kelly P, Doherty A, Mizdrak A, Marshall S, Kerr J, Legge A, et al. High group level validity but high random error of a self-report travel diary, as assessed by wearable cameras. J Transp Heal [Internet]. 2014;1(3):190–201. Available from: http://linkinghub.elsevier.com/retrieve/pii/S2214140514000358

60. Kelly P, Doherty AR, Hamilton A, Matthews A, Batterham AM, Nelson M, et al. Evaluating the Feasibility of Measuring Travel to School Using a Wearable Camera. Am J Prev Med [Internet]. 2012 Nov;43(5):546–50. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0749379712005314

61. Oliver M, Doherty AR, Kelly P, Badland HM, Mavoa S, Shepherd J, et al. Utility of passive photography to objectively audit built environment features of active transport journeys: an observational study. Int J Health Geogr. 2013;12(1):20.

62. Nebeker C, Lagare T, Takemoto M, Lewars B, Crist K, Bloss CS, et al. Engaging research participants to inform the ethical conduct of mobile imaging, pervasive sensing, and location tracking research. Transl Behav Med [Internet]. 2016 Dec 29;6(4):577–86. Available from: https://academic.oup.com/tbm/article/6/4/577-586/4563223

63. Wilson G, Jones D, Schofield P, Martin DJ. Experiences of using a wearable camera to record activity, participation and health-related behaviours: Qualitative reflections of using the Sensecam. Digit Heal [Internet]. 2016 Jan 1;2:205520761668262. Available from: http://journals.sagepub.com/doi/10.1177/2055207616682628

64. Gage R, Leung W, Stanley J, Reeder A, Mackay C, Chambers T, et al. Studying third-parties and environments: New Zealand sun-safety

research. Health Promot Int [Internet]. 2019 Jun 1;34(3):440–6. Available from: https://academic.oup.com/heapro/article/34/3/440/4748797

65. Watkins L, Aitken R, Gage R, Smith MB, Chambers TJ, Barr M, et al. Capturing the Commercial World of Children: The Feasibility of Wearable Cameras to Assess Marketing Exposure. J Consum Aff [Internet]. 2019 Dec 18;53(4):1396–420. Available from: https://onlinelibrary.wiley.com/doi/10.1111/joca.12234

66. Imtiaz MH, Hossain D, Senyurek VY, Belsare P, Tiffany S, Sazonov E. Wearable Egocentric Camera as a Monitoring Tool of Free-Living Cigarette Smoking: A Feasibility Study. Nicotine Tob Res [Internet]. 2020 Oct 8;22(10):1883–90. Available from: https://academic.oup.com/ntr/article/22/10/1883/5613972

67. Zhou Q, Wang D, Mhurchu CN, Gurrin C, Zhou J, Cheng Y, et al. The use of wearable cameras in assessing children's dietary intake and behaviours in China. Appetite [Internet]. 2019 Aug;139:1–7. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0195666318314569

68. Veatupu L, Puloka V, Smith M, McKerchar C, Signal L. Me'akai in Tonga: Exploring the Nature and Context of the Food Tongan Children Eat in Ha'apai Using Wearable Cameras. Int J Environ Res Public Health [Internet]. 2019 May 14;16(10):1681. Available from: https://www.mdpi.com/1660-4601/16/10/1681

69. Laskaris Z, Milando C, Batterman S, Mukherjee B, Basu N, O'neill MS, et al. Derivation of Time-Activity Data Using Wearable Cameras and Measures of Personal Inhalation Exposure among Workers at an Informal Electronic-Waste Recovery Site in Ghana. Ann Work Expo Heal [Internet]. 2019 Oct 11;63(8):829–41. Available from: https://academic.oup.com/annweh/article/63/8/829/5537141

70. Uganda Bureau of Statistics, ICF. Uganda Demographic and Health Survey 2016. Kampala, Uganda and Rockville, Maryland, USA: UBOS and ICF; 2018.

71. World Health Organization. Indicators for assessing infant and young

child feeding practices: Part 1: Definitions. World Health Organization (WHO); 2008.

72. Uganda Bureau of Statistics (UBOS). Time Use Survey 2017/2018 Report. Kampala, Uganda: UBOS; 2019.

73. Kennedy GL, Pedro MR, Seghieri C, Nantel G, Brouwer I. Dietary Diversity Score Is a Useful Indicator of Micronutrient Intake in Non-Breast-Feeding Filipino Children. J Nutr [Internet]. 2007;137(2):472–7. Available from: http://jn.nutrition.org/content/137/2/472

74. Steyn NP, Nel JH, Nantel G, Kennedy G, Labadarios D. Food variety and dietary diversity scores in children: Are they good indicators of dietary adequacy? 2006;9(5):644–50.

75. Hatløy A, Torheim LE, Oshaug A. Food variety—a good indicator of nutritional adequacy of the diet? A case study from an urban area in Mali, West Africa. Eur J Clin Nutr [Internet]. 1998 Dec 17;52(12):891–8. Available from: http://www.nature.com/articles/1600662

76. Moursi MM, Arimond M, Dewey KG, Tre S, Ruel MT, Delpeuch F, et al. Dietary Diversity Is a Good Predictor of the Micronutrient Density of the Diet of 6- to 23-Month-Old Children in Madagascar. J Nutr [Internet]. 2008;138(July):0–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19022971

77. Rah JH, Akhter N, Semba RD, Pee S de, Bloem MW, Campbell AA, et al. Low dietary diversity is a predictor of child stunting in rural Bangladesh. Eur J Clin Nutr [Internet]. 2010;64(12):1393–8. Available from: http://www.nature.com/doifinder/10.1038/ejcn.2010.171

78. Arimond M, Ruel MT. Dietary Diversity Is Associated with Child Nutritional Status: Evidence from 11 Demographic and Health Surveys. J Nutr. 2004;134(10):2579–85.

79. Blackden CM, Wodon Q. Gender, Time Use, and Poverty in Sub-Saharan Africa [Internet]. World Bank Working Papers. Washington, D.C.: The World Bank; 2006. 172 p. Available from: http://elibrary.worldbank.org/doi/book/10.1596/978-0-8213-6561-8

80. Popkin BM. Time allocation of the mother and child nutrition. Ecol Food Nutr. 1980;9(1):1–14.

81. Gryboski KL. Maternal and Non-maternal Time Allocation to Infant Care, and Care during Infant Illness in Rural Java, Indonesia. Soc Sci Med. 1996;43(2):209–19.

82. Komatsu H, Malapit HJL, Theis S. Does women's time in domestic work and agriculture affect women's and children's dietary diversity? Evidence from Bangladesh, Nepal, Cambodia, Ghana, and Mozambique. Food Policy. 2018;

83. Hawkes K, O'Connell JF, Blurton Jones NG. Hadza Women's Time Allocation, Offspring Provisioning, and the Evolution of Long Postmenopausal Life Spans. Curr Anthropol [Internet]. 1997;38(4):551–77. Available from: http://www.journals.uchicago.edu/doi/10.1086/204646

84. Nordang S, Shoo T, Holmboe-Ottesen G, Kinabo J, Wandel M. Women's work in farming, child feeding practices and nutritional status among under-five children in rural Rukwa, Tanzania. Br J Nutr. 2015;114(10):1594–603.

85. van den Bold M, Bliznashka L, Ramani G, Olney D, Quisumbing A, Pedehombga A, et al. Nutrition-sensitive agriculture programme impacts on time use and associations with nutrition outcomes. Matern Child Nutr [Internet]. 2021 Apr 10;17(2):16. Available from: https://doi.org/10.1111/mcn.13104

86. Bamji MS, Thimayamma BVSS. Impact of women's work on maternal and child nutrition. Ecol Food Nutr [Internet]. 2000;39(1):13–31. Available from: http://www.tandfonline.com/doi/abs/10.1080/03670244.2000.9991602

87. Hirway I, Jose S. Understanding Women's Work Using Time-Use Statistics: The Case of India. Fem Econ [Internet]. 2011 Oct;17(4):67–92. Available from:

http://www.tandfonline.com/doi/abs/10.1080/13545701.2011.622289

88. Fedick CB, Pacholok S, Gauthier AH. Methodological issues in the estimation of parental time – Analysis of measures in a Canadian time-use survey. Int J Time Use Res. 2005;2(1):67–87.

89. Buvinić M, King EM. Invisible No More? A Methodology and Policy Review of How Time Use Surveys Measure Unpaid Work [Internet]. United Nations Foundation, Data2X; 2018. Available from: https://data2x.org/resource-center/invisible-no-more-a-methodology-and-policy-review-of-how-time-use-surveys-measure-unpaid-work-volume-1/

90. Walthery P, Gershuny J. Improving Stylised Working Time Estimates with Time Diary Data: A Multi Study Assessment for the UK. Soc Indic Res [Internet]. 2019 Aug 15;144(3):1303–21. Available from: http://link.springer.com/10.1007/s11205-019-02074-3

91. Kanwar P, Yadav D, Sharma N. Time use pattern of hill farm women: A study in Himachal Pradesh. Himachal J Agric Res. 2003;29(1&2):84–8.

92. Szalai A. Women's time: Women in the light of contemporary time-budget research. Futures [Internet]. 1975 Oct;7(5):385–99. Available from: https://linkinghub.elsevier.com/retrieve/pii/0016328775900178

93. Engle P, Lhotská L, Armstrong H, UNICEF. The Care Initiative: Assessment, Analysis and Action to Improve Care for Nutrition [Internet]. UNICEF. New York: UNICEF; 1997. Available from: http://www.sidalc.net/cgi-bin/wxis.exe/?IsisScript=BINCA1.xis&method=post&formato=2&cantidad=1&expresion=mfn=001084

94. Floro MS, King EM. The present and future of time-use analysis in developing countries. Asia-Pacific Popul J [Internet]. 2016 Oct 24;31(1):5–42. Available from: https://www.un-ilibrary.org/population-and-demography/the-present-and-future-of-time-use-analysis-in-developing-countries_748e616d-en

95. Johnston D, Stevano S, Malapit HJL, Hull E, Kadiyala S. Agriculture,

gendered time use, and nutritional outcomes: A systematic review. Washington, D.C.: The International Food Policy Research Institute (IFPRI); 2015. (IFPRI Discussion Paper). Report No.: 01456.

96. Mcguire JS, Popkin BM. Helping Women Improve Nutrition in the Developing World. Washington, D.C.: The International Bank for Reconstruction and Development/THE WORLD BANK; 1990.

97. Floro MS. Economic restructuring, gender and the allocation of time. World Dev [Internet]. 1995 Nov;23(11):1913–29. Available from: https://linkinghub.elsevier.com/retrieve/pii/0305750X9500092Q

98. Nti CA, Inkumsah D, Fleischer G. Influence of women's workload on their nutritional status in selected communities in Ghana. J Consum Stud Home Econ [Internet]. 1999;23(3):165–70. Available from: http://onlinelibrary.wiley.com/doi/10.1046/j.1365-2737.1999.00106.x/abstract

99. Leslie J. Women's time: A factor in the use of child survival technologies? Health Policy Plan [Internet]. 1989;4(1):1–16. Available from: https://academic.oup.com/heapol/article-lookup/doi/10.1093/heapol/4.1.1

100. Leslie J. Women's Work and Child Nutrition in the Third World. In: Women, Work, and Child Welfare in the Third World. 1989.

101. FAO. Minimum dietary diversity for women: An updated guide for measurement. Rome: Food and Agriculture Organization of the United Nations (FAO); 2021.

102. Grosh M, Glewwe P, editors. Designing household survey questionnaires for developing countries: Lessons from 15 years of the Living Standards Measurement Study. Vol. 3. Washington, D.C.: The World Bank; 2000. Report No.: 20731.

103. Alkire S, Meinzen-Dick R, Peterman A, Quisumbing AR, Seymour G. The women's empowerment in agriculture index [Internet]. Oxford: Oxford Poverty & Human Development Initiative; 2013. (OPHI WORKING PAPER). Report No.: 58. Available from: http://www.ophi.org.uk/wp-

content/uploads/ophi-wp-58.pdf?cda6c1

104. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol [Internet]. 2018 Dec 19;18(1):143. Available from: https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-018-0611-x

105. Pearson A, Bottomley R, Chambers T, Thornton L, Stanley J, Smith M, et al. Measuring Blue Space Visibility and 'Blue Recreation' in the Everyday Lives of Children in a Capital City. Int J Environ Res Public Health [Internet]. 2017 May 26;14(6):563. Available from: http://www.mdpi.com/1660-4601/14/6/563

106. Kelly P, Marshall SJ, Badland H, Kerr J, Oliver M, Doherty AR, et al. An Ethical Framework for Automated, Wearable Cameras in Health Behavior Research. Am J Prev Med [Internet]. 2013;44(3):314–9. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0749379712008628

107. Hanley-Cook GT, Tung JYA, Sattamini IF, Marinda PA, Thong K, Zerfu D, et al. Minimum Dietary Diversity for Women of Reproductive Age (MDD-W) Data Collection: Validity of the List-Based and Open Recall Methods as Compared to Weighed Food Record. Nutrients [Internet]. 2020 Jul 9;12(7):2039. Available from: https://www.mdpi.com/2072-6643/12/7/2039

108. European Union. Harmonised European Time Use Surveys (HETUS) 2018 Guidelines. Luxembourg: Publications Office of the European Union; 2019.

109. Bland JM, Altman DG. Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement. Lancet. 1986 Feb;327(8476):307–10.

110. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-

related patient-reported outcomes. J Clin Epidemiol [Internet]. 2010 Jul;63(7):737–45. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0895435610000909

111. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine: a practical guide. Practical guides to biostatistics and epidemiology. Cambridge ; New York: Cambridge University Press; 2011. 338 p.

112. FAO and FHI 360. Minimum Dietary Diversity for Women: A Guide for Measurement. Rome: FAO; 2016.

113. World Health Organization. Indicators for assessing infant and young child feeding practices: Part 2: Measurement. Geneva: World Health Organization (WHO); 2010.

114. Ounpuu S. Seasonality, child nutrition and women's activity patterns: A case study in Chilunga Village, Malawi. [Guelph, Ontario, Canada]: University of Guelph; 1988.

115. Giavarina D. Understanding Bland Altman analysis. Biochem Medica [Internet]. 2015;25(2):141–51. Available from: http://www.biochemia-medica.com/en/journal/25/2/10.11613/BM.2015.015

116. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res. 1999;(8):135–60.

117. Gibson RS. Principles of nutritional assessment. 2nd ed. New York: Oxford University Press; 2005. 908 p.

118. Lesotho B of S. Household Budget Survey Report, Lesotho. Maseru, Lesotho; 2003.

119. Ericsson KA, Simon HA. Verbal reports as data. Psychol Rev [Internet]. 1980;87(3):215–51. Available from: http://content.apa.org/journals/rev/87/3/215

120. Willis GB. Cognitive Interviewing: A "How To" Guide. Research Triangle Park: Research Triangle Institute; 1999.

121. World Health Organization. The WHO Child Growth Standards. Geneva: World Health Organization (WHO); 2006.

122. Schreiner M. Uganda 2012 Poverty Probability Index (PPI) [Internet]. New Haven, Connecticut: Innovations for Poverty Action (IPA); 2016. Available from: www.povertyindex.org/country/uganda

123. Yves Martin-Prevel, Arimond M, Allemand P, Wiesmann D, Ballard TJ, Deitchler M, et al. Development of a Dichotomous Indicator for Population-Level Assessment of Dietary Diversity in Women of Reproductive Age. Curr Dev Nutr [Internet]. 2017;1(12):1–10. Available from: https://academic.oup.com/cdn/article/4772173

124. United Nations Statistics Division. International Classification of Activities for Time-Use Statistics 2016 (ICATUS 2016). New York: United Nations; 2019.

125. Lombard MJ, Steyn NP, Charlton KE, Senekal M. Application and interpretation of multiple statistical tests to evaluate validity of dietary intake assessment methods. Nutr J [Internet]. 2015 Dec 22;14(1):40. Available from: http://nutritionj.biomedcentral.com/articles/10.1186/s12937-015-0027-y

126. Altman DG, Bland JM. Measurement in Medicine: The Analysis of Method Comparison Studies. Stat. 1983;(32):307–17.

127. Ji Y, Plourde H, Bouzo V, Kilgour RD, Cohen TR. Validity and Usability of a Smartphone Image-Based Dietary Assessment App Compared to 3-Day Food Diaries in Assessing Dietary Intake Among Canadian Adults: Randomized Controlled Trial. JMIR mHealth uHealth [Internet]. 2020 Sep 9;8(9):e16953. Available from: https://mhealth.jmir.org/2020/9/e16953

128. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics [Internet]. 1977;33(1):159. Available from: https://www.jstor.org/stable/2529310?origin=crossref

129. Farooq M, Doulah A, Parton J, McCrory M, Higgins J, Sazonov E. Validation of Sensor-Based Food Intake Detection by Multicamera Video Observation in an Unconstrained Environment. Nutrients [Internet].

2019;11(3):609. Available from: https://www.mdpi.com/2072-6643/11/3/609

130. Peterson RA. A Meta-Analysis of Cronbach's Coefficient Alpha. J Consum Res. 1994;21(2):381–391.

131. Bulungu ALS, Palla L, Priebe J, Forsythe L, Katic P, Varley G, et al. Validation of a life-logging wearable camera method and the 24-h diet recall method for assessing maternal and child dietary diversity. Br J Nutr [Internet]. 2021 Jun 14;125(11):1299–309. Available from: https://www.cambridge.org/core/product/identifier/S0007114520003530/type/journal_article

132. Gemming L. Image-assisted dietary assessment: Evaluating the potential of wearable cameras to enhance self -report in the 24-hour dietary recall method. [Auckland, New Zealand]: The University of Auckland; 2015.

133. Gemming L, Ni Mhurchu C. Dietary under-reporting: What foods and which meals are typically under-reported? Eur J Clin Nutr [Internet]. 2016;70(5):640–1. Available from: http://www.nature.com/doifinder/10.1038/ejcn.2015.204

134. Chan V, Davies A, Wellard-Cole L, Lu S, Ng H, Tsoi L, et al. Using Wearable Cameras to Assess Foods and Beverages Omitted in 24 Hour Dietary Recalls and a Text Entry Food Record App. Nutrients [Internet]. 2021 May 26;13(6):1806. Available from: https://www.mdpi.com/2072-6643/13/6/1806

135. Bruun A, Stentoft ML. Lifelogging in the Wild: Participant Experiences of Using Lifelogging as a Research Tool. In: D. Lamas, F. Loizides, L. Nacke, H. Petrie, M. Winckler & PZ, editor. Proceedings of the 17th IFIP TC13 International Conference on Human-Computer Interaction – INTERACT 2019 [Internet]. 2019. p. 431–51. Available from: http://link.springer.com/10.1007/978-3-030-29387-1_24

136. Silva AR, Pinho S, Macedo LM, Moulin CJ. Benefits of SenseCam Review on Neuropsychological Test Performance. Am J Prev Med [Internet]. 2013 Mar;44(3):302–7. Available from:

https://linkinghub.elsevier.com/retrieve/pii/S0749379712008501

137. Hodges S, Williams L, Berry E, Izadi S, Srinivasan J, Butler A, et al. SenseCam: A Retrospective Memory Aid. In 2006. p. 177–93. Available from: http://link.springer.com/10.1007/11853565_11

138. Chow TE, Rissman J. Neurocognitive mechanisms of real-world autobiographical memory retrieval: Insights from studies using wearable camera technology. Ann N Y Acad Sci [Internet]. 2017 May 26;1396(1):202–21. Available from: https://onlinelibrary.wiley.com/doi/10.1111/nyas.13353

139. Small L, Sidora-Arcoleo K, Vaughan L, Creed-Capsel J, Chung K-Y, Stevens C. Validity and Reliability of Photographic Diet Diaries for Assessing Dietary Intake Among Young Children. ICAN Infant, Child, Adolesc Nutr [Internet]. 2009;1(1):27–36. Available from: http://journals.sagepub.com/doi/10.1177/1941406408330360

140. Wilson G, Aitken D, Hodgson P, Bailey C. The hidden impact of home adaptations: Using a wearable camera to explore lived experiences and taken-for-granted behaviours. Health Soc Care Community [Internet]. 2019 Nov 8;27(6):1469–80. Available from: https://onlinelibrary.wiley.com/doi/10.1111/hsc.12818

141. Liu W, Barr M, Pearson AL, Chambers T, Pfeiffer KA, Smith M, et al. Space-time analysis of unhealthy food advertising: New Zealand children's exposure and health policy options. Health Promot Int. 2019;1–9.

142. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet [Internet]. 2007 Oct;370(9596):1453–7. Available from: https://linkinghub.elsevier.com/retrieve/pii/S014067360761602X

143. Bulungu ALS, Palla L, Priebe J, Forsythe L, Katic P, Varley G, et al. Validation of an Automated Wearable Camera-Based Image-Assisted Recall Method and the 24-Hour Recall Method for Assessing Women's

Time Allocation in a Nutritionally Vulnerable Population: The Case of Rural Uganda. Nutrients [Internet]. 2022 Apr 27;14(9):1833. Available from: https://www.mdpi.com/2072-6643/14/9/1833

144. Bulungu ALS, Palla L, Priebe J, Forsythe L, Katic P, Varley G, et al. UNDER REVIEW: Validation of an automated wearable camera-based image-assisted recall method and the 24-hour activity recall method for assessing women's time allocation in rural Uganda. Soc Indic Res. 2022;

145. Doherty AR, Moulin CJA, Smeaton AF. Automatically assisting human memory: A SenseCam browser. Memory [Internet]. 2011 Oct;19(7):785–95. Available from: http://www.tandfonline.com/doi/abs/10.1080/09658211.2010.509732

146. Zuppinger C, Taffé P, Burger G, Badran-Amstutz W, Niemi T, Cornuz C, et al. Performance of the Digital Dietary Assessment Tool MyFoodRepo. Nutrients [Internet]. 2022 Feb 1;14(3):635. Available from: https://www.mdpi.com/2072-6643/14/3/635

147. Jia W, Wu Z, Ren Y, Cao S, Mao Z-H, Sun M. Estimating Dining Plate Size From an Egocentric Image Sequence Without a Fiducial Marker. Front Nutr [Internet]. 2021 Jan 14;7. Available from: https://www.frontiersin.org/articles/10.3389/fnut.2020.519444/full

148. Mohanty SP, Singhal G, Scuccimarra EA, Kebaili D, Héritier H, Boulanger V, et al. The Food Recognition Benchmark: Using Deep Learning to Recognize Food in Images. Front Nutr [Internet]. 2022 May 6;9. Available from: https://www.frontiersin.org/articles/10.3389/fnut.2022.875143/full

149. Gao A, Lo FP-W, Lo B. Food volume estimation for quantifying dietary intake with a wearable camera. In: Body Sensor Networks Conference 2018. IEEE; 2017.

150. Yue Y, Jia W, Fernstrom JD, Sclabassi RJ, Fernstrom MH, Yao N, et al. Food volume estimation using a circular reference in image-based dietary studies. In: Proceedings of the 2010 IEEE 36th Annual Northeast Bioengineering Conference (NEBEC) [Internet]. IEEE; 2010. p. 1–2.

Available from: http://ieeexplore.ieee.org/document/5458155/

151. Sun M, Fernstrom JD, Jia W, Hackworth SA, Yao N, Li Y, et al. A Wearable Electronic System for Objective Dietary Assessment. J Am Diet Assoc [Internet]. 2010;110(1):45–7. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0002822309016897

152. Jia W, Li Y, Qu R, Baranowski T, Burke LE, Zhang H, et al. Automatic food detection in egocentric images using artificial intelligence technology. Public Health Nutr [Internet]. 2019 Mar 26;1–12. Available from: https://www.cambridge.org/core/product/identifier/S1368980018000538/type/journal_article

153. Doherty AR, Caprani N, Conaire CÓ, Kalnikaite V, Gurrin C, Smeaton AF, et al. Passively recognising human activities through lifelogging. Comput Human Behav [Internet]. 2011 Sep;27(5):1948–58. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0747563211000963

154. Doherty AR, Smeaton AF. Automatically Segmenting LifeLog Data into Events. In: 2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services [Internet]. IEEE; 2008. p. 20–3. Available from: http://ieeexplore.ieee.org/document/4556872/

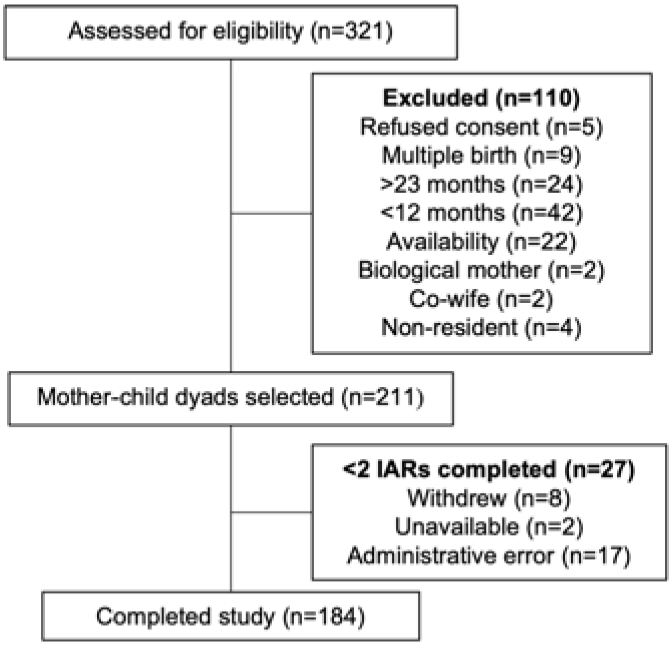
155. Talavera E, Leyva-Vallina M, Sarker MK, Puig D, Petkov N, Radeva P. Hierarchical approach to classify food scenes in egocentric photo-streams. 2018.

156. The World Bank. World Bank Country and Lending Groups [Internet]. [cited 2022 Jun 7]. Available from: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

SUPPLEMENTARY MATERIALS

Supplementary Figure 1. Illustrative photos obtained from the iON SnapCam Lite AWC

Supplementary Figure 2. Study population.



Supplementary Table 1. Characteristics of households and mothers included in and excluded from the study.

	Included				Excluded				*P*
	n (%)	Median	25^{th}-75^{th}		n (%)	Median	25^{th}-75^{th}		
Number of household members		6.0	4.5, 8.0			7.0	5.0, 8.5		.55
Living below $1.25/day (2005 PPP)	149 (20.1)				14 (20.1)				.65
Age (years)		26	22, 30			23	20, 29		.16
Marital status									.51
Married or living together	154 (85.6)				23 (95.8)				
Divorced / Separated or Widowed	8 (4.4)				0 (0.0)				
Never married and never lived together	18 (10.0)				1 (4.2)				
Level of education - attended									1.00
None	13 (7.1)				1 (4.2)				
Primary	125 (67.9)				17 (70.8)				
Post-primary	46 (25.0)				6 (25.0)				
Cannot read and write	90 (50.0)				12 (50.0)				1.00
Christian	108 (58.7)				15 (62.5)				.72

PPP, purchasing power parity; P, p-value using Mann-Whitney U test to compare the medians and Fisher's Exact test to compare the categorical data.

Supplementary Table 2. Participants' rating of their experience with the automated wearable camera-based image-assisted recall method.[a] (N=184)

Rating	n (%)
Good or Very Good	170 (92.4)
Good	67 (36.4)
Very good	103 (56.0)
Bad or Very bad	9 (4.9)
Bad	9 (4.9)
Very bad	0 (0.0)

[a] How would you rate your experience with wearing the camera and looking at the photographs the next day?

Supplementary Table 3a. Participants' most favourite method.

	n (%)	P[1]
AWC-IAR	44 (25.9)	<0.001
OBS / WFR	67 (39.4)	
24HR	7 (4.1)	
MP-IVR	52 (30.6)	

[1] P, p-value using the chi-square goodness-of-fit test to compare the categorical data

Supplementary Table 3b. Participants' most favourite method: AWC-IAR vs OBS / WFR

	n (%)	P[1]
AWC-IAR	44 (25.9)	0.029
OBS / WFR	67 (39.4)	

[1] P, p-value using the chi-square goodness-of-fit test to compare the categorical data.
* Indicates significant at 5% level, with the Bonferroni adjustment for 3 hypothesis (α=0.017).

Supplementary Table 3c. Participants' most favourite method: AWC-IAR vs 24HR

	n (%)	P[1]
AWC-IAR	44 (25.9)	<0.001*
24HR	7 (4.1)	

[1] P, p-value using the chi-square goodness-of-fit test to compare the categorical data
* Indicates significant at 5% level, with the Bonferroni adjustment for 3 hypothesis (α=0.017).

Supplementary Table 3d. Participants' most favourite method: AWC-IAR vs MP-IVR

	n (%)	P[1]
AWC-IAR	44 (25.9)	0.414
MP-IVR	52 (30.6)	

[1] P, p-value using the chi-square goodness-of-fit test to compare the categorical data
* Indicates significant at 5% level, with the Bonferroni adjustment for 3 hypothesis (α=0.017).

Supplementary Table 4a. Participants' least favourite method.

	n (%)	P[1]
AWC-IAR	37 (28.5)	0.041
OBS / WFR	31 (23.9)	
24HR	42 (32.3)	
MP-IVR	20 (15.4)	
[1] P, p-value using the chi-square goodness-of-fit test to compare the categorical data		

Supplementary Table 4b. Participants' least favourite method: AWC-IAR vs OBS / WFR

	n (%)	P[1]
AWC-IAR	37 (28.5)	0.467
OBS / WFR	31 (23.9)	
[1] P, p-value using the chi-square goodness-of-fit test to compare the categorical data. * Indicates significant at 5% level, with the Bonferroni adjustment for 3 hypothesis (α=0.017).		

Supplementary Table 4c. Participants' least favourite method: AWC-IAR vs 24HR

	n (%)	P[1]
AWC-IAR	37 (28.5)	0.574
24HR	42 (32.3)	
[1] P, p-value using the chi-square goodness-of-fit test to compare the categorical data * Indicates significant at 5% level, with the Bonferroni adjustment for 3 hypothesis (α=0.017).		

Supplementary Table 4d. Participants' least favourite method: AWC-IAR vs MP-IVR

	n (%)	P[1]
AWC-IAR	37 (28.5)	0.024
MP-IVR	20 (15.4)	
[1] P, p-value using the chi-square goodness-of-fit test to compare the categorical data * Indicates significant at 5% level, with the Bonferroni adjustment for 3 hypothesis (α=0.017).		

Supplementary Table 5. Insufficient number of images due to AWC inoperability, by week

Issue		Participants Affected by Week							
	Total (N=184)	Bugiri (N=142)					Kamuli (N=69)		
		WK1	*WK2*	*WK3*	*WK4*	*LD Total*	*WK5*	*WK6*	*LD Total*
	n (%)	*n*	*n*	*n*	*n*	*n (%)*	*n*	*n*	*n (%)*
Participants with insufficient number of images due to AWC inoperability*	27 (14.7)	0	4	1	4	9 (6.3)	7	11	18 (26.1)

* Defined as 12 hours of images (1,440) per 15-hour day (1,800).
WK, week; IAR, image-assisted recall.

Supplementary Table 6. Participant's ability to recognize selected image types during the first and second image-assisted recall orientation.

Orientation photo type	IAR1	IAR2	*P**
	n (%)	n (%)	
Photo of herself			0.688
Recognized	171 (98.8)	172 (97.7)	
Did not recognize	2 (1.2)	4 (2.3)	
Photo of her child			0.500
Recognized	184 (100.0)	180 (98.9)	
Did not recognize	0 (0.0)	2 (1.1)	
Photo of her home			1.000
Recognized	180 (97.8)	178 (97.3)	
Did not recognize	4 (2.2)	5 (2.7)	
Photo of her garden			0.039
Recognized	136 (93.2)	142 (96.6)	
Did not recognize	10 (6.9)	5 (3.4)	
Photo of her own hand while performing a task			0.036
Recognized	164 (89.1)	175 (95.6)	
Did not recognize	20 (10.9)	8 (4.4)	

P, p-value using McNemar's exact test to compare the paired categorical data.

Supplementary Table 7. Lost data due to AWC inoperability - with and without observer present

	Lost data due to inoperability	
	Observer present	Observer NOT present
AWC-IAR1 (n=184)	8	20
AWC-IAR2 (n=184)	16	12
Total	24 (6.5)	32 (8.7)

ADDITIONAL MATERIALS (NOT INCLUDED IN THE MANUSCRIPT)

Additional Table 1. Characteristics of households, mothers, and children participating in and lost to the study.

	Participating		Missing		*P*
	n	%	n	%	
Mothers					
Mobile phone type					0.870
Button phone	156	96.9	26	96.3	
Smart phone	5	3.1	1	3.7	
Frequency of mobile phone use					0.333
Daily	46	31.3	10	37.0	
Weekly	47	32.0	10	37.0	
Rarely	54	36.7	7	25.9	

P, p-value using Mann-Whitney U test to compare the medians and Fisher's Exact test to compare the categorical data.

Additional characteristics of the study population are presented and compared with participants who were lost to follow-up or excluded from the analyses, in Additional Table 1. Most participants used button phones. The frequency of mobile phone use varied but was equally distributed among participants. There were no significant differences between participants and those lost to follow-up or excluded from the study.

Notably, only 3% of participants had access to a smart phone and, although no other data was collected in this study regarding access to other visual media, in 2016 only 7% of rural Ugandans owned a television. (70) Participant's exposure to first-person perspective imagery may therefore be limited.

Chapter 5: Validation of a life-logging wearable camera method and the 24-h diet recall method for assessing maternal and child dietary diversity

ABSTRACT

Accurate and timely data are essential for identifying populations at risk for undernutrition due to poor-quality diets, for implementing appropriate interventions and for evaluating change. Life-logging wearable cameras (LLWC) have been used to prospectively capture food/beverage consumed by adults in HICs. This study aimed to evaluate the concurrent criterion validity, for assessing maternal and child dietary diversity scores (DDS), of a LLWC-based image-assisted recall (IAR) and 24-h recall (24HR). Direct observation was the criterion method. Food/beverage consumption of rural Eastern Ugandan mothers and their 12–23-month-old child (n 211) was assessed, for the same day for each method, and the IAR and 24HR DDS were compared with the weighed food record DDS using the Bland–Altman limits of agreement (LOA) method of analysis and Cohen's κ. The relative bias was low for the 24HR (–0·1801 for mothers; –0·1358 for children) and the IAR (0·1227 for mothers; 0·1104 for children), but the LOA were wide (–1·6615 to 1·3012 and –1·6883 to 1·4167 for mothers and children via 24HR, respectively; –2·1322 to 1·8868 and –1·7130 to 1·4921 for mothers and children via IAR, respectively). Cohen's κ, for DDS via 24HR and IAR, was 0·68 and 0·59, respectively, for mothers, and 0·60 and 0·59, respectively, for children. Both the 24HR and IAR provide an accurate estimate of median dietary diversity, for mothers and their young child, but non-differential measurement error would attenuate associations between DDS and outcomes, thereby under-estimating the true associations between DDS – where estimated via 24HR or IAR – and outcomes measured.

RESEARCH PAPER COVER SHEET

LONDON SCHOOL of HYGIENE & TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1407282	**Title**	Ms.
First Name(s)	Andrea		
Surname/Family Name	Spray		
Thesis Title	Validation of an automated wearable camera-based image-assisted recall method and the 24-hour recall method for assessing maternal and child dietary diversity and women's time use in rural Uganda		
Primary Supervisor	Elaine Ferguson		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	British Journal of Nutrition		
When was the work published?	2021		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Improving health worldwide **www.lshtm.ac.uk**

Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	A. L. S. B. contributed to the study design and tool development, coordinated data collection and contributed to study management, developed data entry tools, entered data and oversaw data entry research assistants, carried out the validation analyses, interpreted the results, and wrote and edited the manuscript.

SECTION E

Student Signature	[redacted]
Date	27/6/2022

Supervisor Signature	[redacted]
Date	15/6/2022

MANUSCRIPT

British Journal of Nutrition (2021), 125, 1299–1309 doi:10.1017/S0007114520003530

Validation of a life-logging wearable camera method and the 24-h diet recall method for assessing maternal and child dietary diversity

Andrea L. S. Bulungu[1]*, Luigi Palla[2], Jan Priebe[3], Lora Forsythe[3], Pamela Katic[3], Gwen Varley[3], Bernice D. Galinda[1], Nakimuli Sarah[1], Joweria Nambooze[4], Kate Wellard[3] and Elaine L. Ferguson[1]

[1]Department of Population Health, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK
[2]Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK
[3]Natural Resources Institute (NRI), University of Greenwich, Chatham Maritime ME4 4TB, UK
[4]Africa Innovations Institute (AfrII), Kampala, Uganda

(Submitted 20 April 2020 – Final revision received 30 July 2020 – Accepted 26 August 2020 – First published online 11 September 2020)

Abstract

Accurate and timely data are essential for identifying populations at risk for undernutrition due to poor-quality diets, for implementing appropriate interventions and for evaluating change. Life-logging wearable cameras (LLWC) have been used to prospectively capture food/beverage consumed by adults in high-income countries. This study aimed to evaluate the concurrent criterion validity, for assessing maternal and child dietary diversity scores (DDS), of a LLWC-based image-assisted recall (IAR) and 24-h recall (24HR). Direct observation was the criterion method. Food/beverage consumption of rural Eastern Ugandan mothers and their 12–23-month-old child (n 211) was assessed, for the same day for each method, and the IAR and 24HR DDS were compared with the weighed food record DDS using the Bland–Altman limits of agreement (LOA) method of analysis and Cohen's κ. The relative bias was low for the 24HR (–0·1801 for mothers; –0·1358 for children) and the IAR (0·1227 for mothers; 0·1104 for children), but the LOA were wide (–1·6615 to 1·3012 and –1·6883 to 1·4167 for mothers and children via 24HR, respectively; –2·1322 to 1·8868 and –1·7130 to 1·4921 for mothers and children via IAR, respectively). Cohen's κ, for DDS via 24HR and IAR, was 0·68 and 0·59, respectively, for mothers, and 0·60 and 0·59, respectively, for children. Both the 24HR and IAR provide an accurate estimate of median dietary diversity, for mothers and their young child, but non-differential measurement error would attenuate associations between DDS and outcomes, thereby under-estimating the true associations between DDS – where estimated via 24HR or IAR – and outcomes measured.

Key words: Validation studies: Diet assessment: Wearable cameras: Bland–Altman method

Globally, undernutrition is the single biggest contributor to child mortality[1]. Although the underlying determinants of undernutrition are complex and interacting, inadequate nutrient intake is an immediate cause[2,3]. The prevalence of undernutrition is decreasing; however, there are still 144 million children under 5 years who suffer from chronic malnutrition, more than a third of whom live in Africa[4]. Accurate and timely data are essential for identifying populations at risk for undernutrition due to poor quality diets, for implementing appropriate interventions and for evaluating change[5–8].

Current methods of quantitative dietary assessment are reliable but resource-intensive. To address the need for a rapid, inexpensive and simple-to-administer method with a low participant burden, reliable population-level food group indicators for measuring diet quality requiring only semi-quantitative dietary data were developed. Cross-country analyses assessing their performance for predicting nutrient adequacy have shown moderate, but variable, associations[9–11]. Two indicators in particular – a dietary diversity score (DDS) for children under 5 years and a DDS for women of reproductive age living in low- and middle-income countries – have been validated and are now in widespread use globally[12–14]. Corresponding global minimum dietary diversity (MDD) standard thresholds have also been validated and are in widespread use[12,13,15]. Assessing dietary diversity using the DDS or other food group-based indicators is, in comparison, much simpler and requires only assessing whether any representative foods of each food group in the index were consumed.

'Gold standard' quantitative dietary assessment methods, such as direct observation and repeated weighed food records (WFR), are used to accurately and reliably assess the foods consumed as well as the amount of eatable portion of each food consumed by individual consumers. They are seldom used to routinely collect dietary data due to high financial and time costs,

Abbreviations: 24HR, 24-h recall; DDS, dietary diversity score; EII, enumerator image interpretation; IAR, image-assisted recall; LIC, low-income country; LLWC, life-logging wearable camera; LOA, limit of agreement; MDD, minimum dietary diversity; WFR, weighed food record.

* Corresponding author: Andrea L. S. Bulungu, email andrealspray@gmail.com

British Journal of Nutrition

participant burden and requisite expertise[16–18]. Instead, retrospective diet assessment tools using a multiple-pass 24-h 'free' recall technique, which entails mothers recalling all foods consumed by their child or themselves on the previous day, are commonly used for estimating food consumption. The accuracy of all food recall methods relies upon the respondent's memory and motivation, as well as the skill and persistence of the interviewer. Misreporting of foods consumed may occur, either unintentionally – for example, for foods that are infrequently consumed – or intentionally, due to interviewer, social desirability or approval biases[14]. Such errors can result in either underreporting or over-reporting (or both) of food groups defining the DDS. The 24-h recall (24HR) method is also susceptible to other measurement errors[19,20].

There is a long history of using photos to overcome the limitations of traditional food recall methods. 'Portion-size estimation aids' involve the use of gradated food photos (representing the range of portion sizes commonly consumed) when collecting semi-quantitative data[21–31]. They are among the earliest approaches to incorporating food photography into diet assessment methods to reduce recall bias and have now been validated in both high- and low-income countries. Another approach to incorporating food photography (or illustrations) into traditional diet assessment methods is the use of a pictorial food chart to help low-literacy subjects in Sub-Saharan Africa prospectively self-record foods consumed[17,32,33].

In high-income country contexts, validation research has also been conducted, in controlled settings, into the use of prospective photos taken by the study subject (often using a mobile phone) during food preparation and/or eating episodes[34–37]. In these studies, information documented in the photos regarding foods consumed, portion sizes and wastage is later used by researchers to compute nutrient intake. Such 'active' photography methods may reduce unintentional errors but do not eliminate the possibility of intentional over- or under-estimation due to social desirability bias. And, although often preferred to traditional methods, active prospective food photo methods require a high level of technical competence among study participants, which further limits its use in large-scale surveys.

New digital media technologies offer opportunities for improving upon traditional 24HR methods of assessing dietary diversity[38–41]. A life-logging wearable camera (LLWC) worn by a study participant can be used to prospectively and passively capture food/beverage consumed, which may provide a more objective method of data collection for assessing women's and children's dietary diversity than the 24HR method, with low respondent and interviewer burden. In high-income countries, limited validation research on image-assisted recall (IAR) methods for assessing quantitative dietary data suggests that a LLWC can reduce underreporting of energy intake[42,43]. Yet, questions remain about their validity, acceptability and feasibility[44,45].

In addition to a range of socio-cultural factors that may affect the acceptability of the LLWC in rural low-income country settings, the contexts pose unique technical challenges for wearable cameras, such as lack of electricity with which to illuminate food preparation and consumption at night or indoors and to charge the wearable devices; rugged conditions that expose the camera to water and dirt; and lack of familiarity among participants to digital technology and social media – in particular to first-person photos – which may hamper their interpretation. Furthermore, logistical challenges can be anticipated in assessing dietary diversity for young children. For example, using a LLWC attached to the caregiver may not fully capture a child's food intake if the child spends substantial time out of the direct supervision of the primary caregiver (and therefore not in sight of the camera).

This study was undertaken in rural Eastern Uganda to evaluate the concurrent criterion validity, for assessing maternal and child DDS and MDD, of a LLWC-based IAR method and the 24HR method. Direct observation with WFR was the criterion method. Previous studies have evaluated the validity of photo-assisted methods to assess nutrient intake in high-income country contexts. No study, to our knowledge, has examined the validity, for estimating DDS or MDD, of the 24HR method or an IAR method using a LLWC in either free-living or controlled settings.

Methods

Study design

A cross-sectional study of mothers and their child aged 12–23 months (*n* 211) was conducted between January and February (dry season) 2018 in Bugiri and Kamuli Districts, Eastern Region, Uganda. This study was nested within another study designed to examine the impact of a labour-saving technology (a mechanised maize sheller) on women's time for childcare, food preparation and dietary practices.

In our study, food/beverage consumption of mothers and their child was assessed, for the same day, using three concurrent methods: (1) direct observation (15 h) via WFR, (2) 24HR and (3) IAR using a LLWC (wearable camera). Data were collected over five consecutive days, following one of the two possible patterns (Fig. 1). Specifically, for both patterns, on day 1, eligibility was confirmed, a structured questionnaire was administered and anthropometric data were collected for all participants. Day 1 data were collected at a predefined meeting place in the village. All other data were collected at the participants' home. For half of the study participants, on day 2, food/beverage consumption data were collected using direct observation via WFR and recorded on the wearable camera attached to the mother. On day 3, a 24HR was administered, followed by an IAR using photos captured on day 2 by the wearable camera. On day 4, food/beverage consumption data were again recorded via the wearable camera only (i.e. no observation). On day 5, an IAR was administered using photos captured on day 4 by the wearable camera. The other half of the study participants began with

HHP	Day 1	Day 2	Day 3	Day 4	Day 5
1	ICF MQ Anthro	O and I	24HR and IAR	I	IAR TQ
2	ICF MQ Anthro	I	IAR	O and I	24HR and IAR TQ

Fig. 1. Data collection household pattern (HHP).

the wearable camera only (i.e. days 2 and 4 were switched) and ended with all three methods (i.e. days 3 and 5 were switched). For all participants, on the 5th day, a final structured questionnaire was also administered. All data collection was performed by trained enumerators. Dietary data collection was distributed across all days of the week to minimise any day-of-the-week effect, and for each mother–child dyad, the enumerator assigned to conduct the direct observation was different from the enumerator assigned to administer the 24HR and IAR.

Ethical approval was obtained from the Uganda National Council for Science and Technology (A24ES), the London School of Hygiene & Tropical Medicine Observational/Interventions Research Ethics Committee (project ID: 1420) and the University of Greenwich Faculty of Engineering and Science Ethics Committee (project ID: B0501). The data collection protocols followed the ethical guidelines for LLWC research to ensure that privacy of the participants was maintained[38]. Following community sensitisation, verbal explanation of the study and demonstration of the wearable camera, written consent or thumb print was obtained from all mothers who participated in our study.

Participants and sampling

Twenty-two villages were purposefully selected, for this study, of which eleven had access to labour-saving technology and eleven did not. These villages participated in the Sasakawa Global 2000 Uganda (SG2000 Uganda) country programme (the local implementing partner for the parent study). The sample size calculation (n 264; twenty-two communities, twelve households per community) was based on requirements of the time allocation study within which this current study was nested. A sample size of 132 per group enabled detection of a 30-min inter-group difference between women with access to a labour-saving device and households without access to a labour-saving device, assuming a SD of 49 min, a design effect of 1·47, 80 % power and P value of 0·05 and allowing for 10 % attrition. This sample size was deemed sufficient for the current validation study, using the Bland–Altman method of analysis[46–49].

The sampling frame, for each village, was a household listing of all mothers with children born between 1 January 2016 and 1 May 2017 inclusive (to recruit children aged 12–24 months at the time of data collection). These lists were generated by the SG2000 community-based facilitators. Twelve mother–child dyads in each village were randomly selected to participate in the study; substitutions were made, as needed, until twelve mother–child dyads who met the inclusion/exclusion criteria were recruited. Mother–child dyads were excluded if the child was <12 months or >23 months of age, was not yet eating solid foods on a regular basis, was a multiple-birth child, the mother was unable to communicate in Lusoga, Luganda or English; either the mother or child had a severe disability; the mother was not the biological mother of the child; the mother was a co-wife with another mother selected to participate in the study; or either the mother or child was not available for the duration of the study.

Instruments and protocol

The enumerators administered two structured questionnaires to the mother. The first questionnaire collected information on: household socio-demographics, wealth (adapted from the Uganda 2012 Poverty Probability Index), expenditure and production (adapted from the Abbreviated Women's Empowerment in Agriculture Index); knowledge, attitudes and practices regarding infant and young child feeding and care; factors related to women's empowerment (adapted from the Abbreviated Women's Empowerment in Agriculture Index) and the child's health (adapted from Demographic and Health Surveys). The second questionnaire was administered at the end of the period of data collection. Response options in this questionnaire included a four-point Likert for questions related to participants' perceptions of their experiences with each of the three food/beverage data collection methods, and categorical scales for questions related to mobile phone access, ownership and use, and willingness to participate in a future study using 24HR or wearable cameras, in addition to open-ended comments. Duplicate, serial anthropometric measurements of weight and height/length were taken from participating mothers and children.

For the criterion diet assessment method (i.e. WFR), enumerators weighed and recorded all food/beverage consumed by mothers and children, from approximately 06.00 to 21.00 hours using dietary scales (±1 g, Salter Disc Electronic Digital Scale Model 1036) and a standard WFR protocol[16]. Recipe data were also collected by weighing the recipe ingredients and final cooked food and recording the cooking methods (e.g. fried, boiled and stewed). If the child was left in the care of another person, the enumerator remained with the mother and information about any foods or beverages consumed while the child was away was collected from the secondary caregiver via recall upon their return. The amounts of food/beverage consumed by the mother or child before 06.00 or after 21.00 hours were recalled and recorded.

On the day after the WFR was collected, two semi-quantitative multiple-pass 24-h dietary recalls were administered to the mother to collect information on all foods and beverages consumed the previous day – one for herself and then one for the child[16]. For each recall, in the first pass, the mother was asked to list everything she (or her child) consumed the previous day; in the second pass, additional details about each food were recorded, including the time of consumption and ingredients in mixed dishes. In the third pass, mothers were asked to confirm the food groups consumed. The quantity of each food consumed was not recorded.

The same day as WFR data collection, a small, lightweight, LLWC was attached to a t-shirt worn by the mother at approximately 06.00 hours and removed at approximately 21.00 hours. Participants were instructed to wear the camera while continuing their usual activities, covering or removing the camera as needed for privacy. The wearable camera automatically recorded a picture every 30 s, storing all photos (approximately 1800) on a memory card.

The following day, an enumerator first reviewed the photos captured by the wearable camera on a tablet and annotated the foods she thought – based on the photos – were consumed by the mother and child, that is, the enumerator image interpretation (EII). The enumerator estimated the DDS for the mother and child based on her interpretation of the photos and demarcated the series of eating episodes for later review with the mother. Upon meeting with the mother, the enumerator first administered the two standard 24HR (one for the mother and one for her child). Then, the enumerator reviewed the photos with the mother on the tablet. During this interview, the enumerator probed the participant based on 'the 4 Ws, where appropriate; for example, questions such as: 'What were you doing?', 'Who were you with?', 'Where were you?', 'Where were you going?', 'Where was the index child?', 'Why did you go there?', or 'Why were you doing that?' The enumerator revised her original annotations (i.e. the EII) of foods consumed by the mother and child, as needed, based on the mother's feedback. Finally, having reviewed and discussed the previous day's photos with the mother, the enumerator asked the mother to confirm the food groups that she and her child had or had not consumed (i.e. the IAR).

The IAR protocol was adapted from one used in high-income country contexts[39]. The protocol followed ethical guidelines for LLWC research to ensure that the privacy of the participants was maintained[38]. The IAR protocol was pilot-tested prior to the start of the study.

Data processing

The food/beverage recorded over a period of 24 h were coded into food groups, and a DDS was calculated for each mother and child for each method (WFR, 24HR and IAR). The DDS for children was based on seven food groups: namely grains, roots and tubers; legumes and nuts; dairy products; flesh foods; eggs; vitamin A-rich fruits and vegetables; and other fruits and vegetables[12]. The DDS for women was based on ten food groups: namely grains, white roots and tubers, and plantains; pulses; nuts and seeds; dairy products; meat, poultry and fish; eggs; dark green leafy vegetables; other vitamin A-rich fruits and vegetables; other vegetables; other fruits[13]. The percentage of women and children achieving MDD was calculated. The threshold MDD used for women was five food groups out of ten; the threshold MDD for children was four food groups out of seven. Breast milk intake was not included in the comparison of the three methods.

Weight-for-age *z*-score, length-for-age *z*-score and weight-for-length *z*-score were calculated for each child using the 2006 WHO growth standards[50], and BMI was calculated for each mother. The proportions of children who were underweight (<–2 SD from median weight-for-age *z*-score), stunted (<–2 SD from median length-for-age *z*-score) and wasted (<–2 SD from median weight-for-length *z*-score) were calculated. The proportions of mothers who were thin (BMI < 18·5 kg/m^2), normal (BMI 18·5–24·9 kg/m^2) and overweight/obese (BMI ≥ 25·0 kg/m^2) were also calculated. The Ugandan 2012 Poverty Probability Index was calculated, as well as the proportion of the population living below $1·25/d[51].

Data analysis

The primary outcome variables analysed for both mothers and children were DDS and MDD. Data were analysed using Stata/SE version 15.1. *P* values <0·05 were considered significant for all tests. Cases with incomplete data for any of the three methods (WFR, 24HR or IAR) were eliminated from analysis. A minimum threshold of 13 h of observation and photos from the wearable camera were deemed adequate; any cases not meeting this threshold were eliminated from analysis to ensure the integrity of the comparison (with the 24HR). The Wilcoxon signed-rank sum and McNemar's tests were used to compare the distributions of DDS and MDD, respectively, obtained via the criterion method (WFR) *v*. the 24HR or IAR. The medians of the DDS differences (24HR minus WFR and IAR minus WFR) were computed, and the distribution of the median DDS differences was also compared. Key socio-demographic characteristics for participating and missing households were compared using the Mann–Whitney two-sample statistic for continuous data, and the Fisher Exact test for categorical data. DDS and MDD for mothers and children in participating households collected via IAR were also calculated for the non-observation day and compared with those of the corresponding observation day, using the Wilcoxon signed-rank sum and McNemar's tests, respectively.

Treating DDS as a continuous measure, the inter-tool agreement between WFR and 24HR or IAR was assessed using the Bland–Altman limits of agreement (LOA) method[46]. Specifically, for each individual, the difference between the methods (DDS estimated using either the 24HR or IAR minus the criterion measure of DDS) *v*. the mean of the methods was plotted; the relative bias and the 95 % LOA (mean difference ± 2 SD of the differences) were estimated. Finally, DDS estimates via the 24HR and IAR methods against the criterion method were also compared using the weighted Cohen's κ coefficient for inter-rater agreement. It was interpreted as follows: <0·00 poor agreement; 0·00–0·20 slight agreement; 0·21–0·40 fair agreement; 0·41–0·60 moderate agreement; 0·61–0·80 substantial agreement; 0·81–1·00 almost perfect agreement[52,53].

Results

Characteristics of the sample

Overall, 211 mother–child dyads were recruited into the study. Among those recruited, six participants voluntarily withdrew and forty-two participants were eliminated from analysis due to incomplete data (Fig. 2). Characteristics of the study population are presented and compared with participants who were lost to the study in Table 1. These comparisons show there were no differences between participating and missing households, with the exception of child breast-feeding status (61 % for participating children *v*. 42 % for non-participating children). The median household size was six members, and nearly one-quarter of participating households lived below $1·25/d. Most participating mothers were married and between the ages of 20 and 29 years. Nearly two-thirds of participating mothers had not completed primary school, and just under one-half were

British Journal of Nutrition

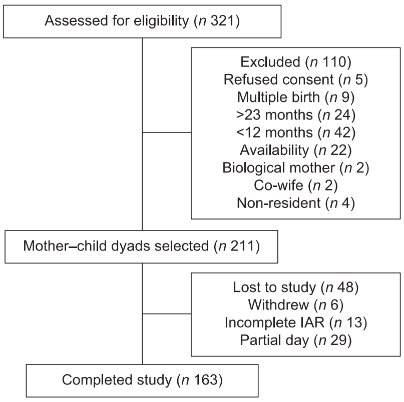


Fig. 2. Study population. IAR, image-assisted recall.

literate. Most mothers were either pregnant, breast-feeding or both. Most women were of normal BMI.

The median age of participating children was 16·7 months, approximately evenly split between males and females. Nearly all children were initially breastfed, although just 61 % were breast-feeding at the time of data collection. Among this population, children were fed by several caregivers in addition to their mother. More than a third of children were fed by at least one caregiver <13 years of age. Approximately a quarter of children were stunted but <3 % of children were wasted.

Diet diversity

The median DDS, for both mothers and children and all methods, was four food groups (Table 2). The estimated percentage of mothers achieving the MDD ranged from 41 % for the WFR to 47 % for the 24HR and for children it ranged from 55 % for the WFR to 60 % for the 24HR (Table 2). The percentage achieving the MDD estimated via the 24HR and IAR, for both mothers and children, was consistently higher than the WFR estimates.

Median DDS and MDD for mothers estimated via IAR on the non-observation day were slightly higher than those collected via IAR on the observation day (5 *v.* 4, *P* = 0·2862; and 54 *v.* 42 %, *P* = 0·1161) (see online Supplementary Tables S1a and b). For children, median DDS and MDD estimated via IAR on the non-observation day were similar to those collected via IAR on the observation day (4, *P* = 0·5243; and 56 *v.* 58 %, *P* = 0·3428) (see online Supplementary Table S1a and b).

Measure of agreement

The Bland–Altman plots showed a consistent and uniform pattern across the range of mean DDS for all analyses (Fig. 3). The relative bias was low for the 24HR (–0·1801 for mothers; –0·1358 for children) and the IAR (–0·1227 for mothers; –0·1104 for children) (Table 3). The percentage of DDS that were identical comparing the IAR or 24HR with the criterion method ranged from 58 % (IAR for mothers) to 70 % (24HR children). Between 6 and 9 % of the estimates were erred by two or more food groups (see online Supplementary Fig. S1). Although the relative bias was not clinically important, the LOA were wide (–1·6615 to 1·3012 and –1·6883 to 1·4167 for mothers and children via 24HR, respectively; –2·1322 to 1·8868 and –1·7130 to 1·4921 for mothers and children via IAR, respectively). For DDS estimated via 24HR and IAR, Cohen's κ coefficient was 0·68 and 0·59, respectively, for the mothers, and 0·60 and 0·59, respectively, for the children (Table 4). For mothers, Cohen's κ indicated slightly higher inter-method agreement for the 24HR (substantial reliability) than the IAR (moderate reliability), whereas for children, the inter-method agreements for the 24HR and IAR are both moderate.

Discussion

This is the first study, to our knowledge, validating the 24HR and IAR method using a LLWC for assessing DDS and MDD. Both the 24HR and the IAR provided an accurate estimate of the sample median DDS for women and young children, for the same day of food intake, but tended to overestimate the proportion of women or children that achieved MDD, indicating that the 24HR and the IAR may over-estimate diet quality, at least among women and young children of Eastern Uganda during the dry season.

Although the relative bias seen in this study was low, the high LOA observed for both methods (24HR and IAR) across population groups (mothers and children) was substantial. There are no validation studies of IAR methods for estimating DDS with which to compare these results. However, similar results have been seen in image-assisted quantitative diet recall validation studies[29,33,35,37,54]. Such error usually serves to attenuate the association between DDS and health or other outcomes and diminish power to detect it[18]. This study indicates that attenuation levels might remain high even after accounting for day-to-day variation in DDS because food groups are misreported by a substantial proportion of individuals.

It is well documented that, in quantitative dietary assessment, measurement error (the difference between reported intake and true intake) commonly occurs[18]. This study suggests that a high degree of measurement error also occurs when diet quality is assessed by the number of selected food groups consumed. In this study, individual estimates of DDS could differ by more than two food groups from observed values. It is remarkable that both 24HR and IAR mis-classified over a third of maternal and child DDS. Although there is no globally accepted threshold LOA for DDS, a difference of one or two food groups (out of seven for children and ten for women) is substantial. Errors appear to occur equally at lower and higher ends of the DDS spectrum.

The cause of the wide LOA observed in this study is not immediately evident. Further exploration of the data shows that reported consumption of vitamin A-rich fruits and vegetables was higher (≥5 %) in the IAR and 24HR than the WFR for mothers and children, and a higher percentage of mothers reported consuming other fruits (62 *v.* 56 %) and dairy products (22 *v.* 17 %) in the IAR compared with WFR (see online Supplementary Table

Table 1. Characteristics of households, mothers and children participating in and lost to the study
(Numbers and percentages; median values and 25th and 75th percentiles)

	Participating				Missing				
	n	%	Median	25th, 75th (%)	*n*	%	Median	25th, 75th (%)	*P*
Households									
Number of household members			6	5, 8			5·5	4, 7·5	0·13
Living below $1·25/d (2005 PPP)			20·1	10·9, 27·9			20·1	4, 45·3	0·163
Mothers									
Age (years)			26	22, 30			23	20, 30	0·088
15–19	17	10·4			7	14·9			0·766
20–29	100	61			28	59·6			
30–39	39	23·8			11	23·4			
40–49	8	4·9			1	2·1			
Marital status									
Single	20	12·3			1	2·3			0·081
Married or co-habiting	135	82·8			42	95·5			
Level of education									
None or primary incomplete	102	62·2			22	50			0·353
Primary complete	55	33·5			19	43·2			
Secondary complete	5	3·1			2	4·6			
Can read and write	78	48·5			24	55·8			0·493
Maternity status									
Pregnant	22	13·9			11	25			0·105
Breast-feeding	104	63·4			23	48·9			0·091
Pregnant or breast-feeding	122	74·4			30	63·8			0·196
BMI			21·7	20·4, 23·9			21·6	19·0, 23·3	0·454
% Thin (BMI < 18·5 kg/m^2)	7	5·4			4	13·8			0·286
% Overweight/obese (BMI ≥ 25·0 kg/m^2)	17	13·1			3	10·3			
Children									
Age (months)			16·7	14·9, 20·0			17·5	14·7, 19·6	0·682
12–17	99	60·7			24	52·2			0·313
18–23	64	39·3			22	47·8			
Sex									
Female	75	45·7			23	50			0·62
Male	89	54·3			23	50			
Ever breastfed	161	99·4			42	97·7			0·376
Currently breast-feeding	98	60·5			18	41·9			0·037
Child feeders			3	2, 4			3	2, 4	0·67
No alternative feeders	19	11·6			7	14·9			0·615
All child feeders > 13 years	101	61·6			32	68·1			0·494
Length-for-age *z*-score			−1·3	−2·1, −0·6			−1·1	−2·2, −0·6	0·939
% below −2 SD	42	25·8			13	29·6			0·508
Weight-for-age *z*-score			−0·7	−1·5, −0·1			−0·7	−1·5, −0·1	0·848
% below −2 SD	14	9			7	18			0·067
Weight-for-length *z*-score			−0·2	−0·8, 0·4			−0·2	−0·6, 0·3	0·776
% below −2 SD	4	2·6			0	0			0·678

PPP, purchasing power parity; *P*, *P* value using Mann–Whitney *U* test to compare the medians and Fisher's exact test to compare the categorical data.

S2a and b), whereas a lower percentage of mothers reported consumption of other vegetables in the IAR compared with the WFR (90 *v.* 95 %).

Our finding that the LOA in assessing DDS were high, for both the 24HR and IAR collected for the same day as the criterion method, was somewhat surprising. Compared with the 24HR, we had expected viewing one's own passively collected photographs (IAR) would reduce errors due to memory, social desirability and other biases commonly known to contribute to inaccurate estimates when recalling foods consumed.

There are several plausible contributing factors. For instance, mothers may have become bored and/or fatigued after four similar series of questions about the food groups consumed, thus resulting in more random error in IAR than 24HR, because the mother and child 24HR were administered before the mother and child IAR. Further review by the author (A. L. S. B.) of the data inconsistencies in the procession from WFR to EII to IAR among participants with large discrepancies in DDS suggests that inflation of the DDS from WFR to the IAR may have been due to errors introduced by the mother during the last step of the IAR protocol (i.e. the final confirmation of food groups consumed). In addition to boredom or fatigue, social desirability bias may have contributed to inflation of DDS. No pattern was observed for those that under-estimated the DDS.

Alternatively, and consistent with the long-standing theory that low levels of education and pictorial literacy may affect subjects' capacity to interpret food photos, some women in this study struggled to interpret the first-person photos from the wearable camera[55]. Enumerators were instructed to record foods consumed as reported by the mother, even if it conflicted

British Journal of Nutrition

Table 2. Inter-method comparisons of the median dietary diversity scores (DDS) and percentage achieving minimum dietary diversity (MDD)
(Median values and 25th and 75th percentiles; numbers and percentages)

	Mother						Child					
	WFR (*n* 163)		24HR (*n* 161)		IAR (*n* 163)		WFR (*n* 163)		24HR (*n* 162)		IAR (*n* 163)	
	Median	25th, 75th (%)	Median	25th, 75th (%)	Median	25th, 75th (%)	Median	25th, 75th (%)	Median	25th, 75th (%)	Median	25th, 75th (%)
DDS*	4	3, 5	4‡	4, 5	4	4, 5	4	3, 4	4‡	3, 4	4	3, 4
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
MDD†	67	41·1	76	47·2§	69	42·3	90	55·2	97	59·9	95	58·3

WFR, weighed food record; 24HR, 24-h recall; IAR, image-assisted recall.
* Mother DDS is out of ten food groups; child DDS is out of seven food groups.
† Mother MDD is DDS ≥ 5; child MDD is DDS ≥ 4.
‡ *P* value of Wilcoxon signed-rank sum test of DDS compared with WFR < 0·05.
§ *P* value of McNemar's test of the MDD proportion difference compared with WFR < 0·05.

with their own interpretation of the photos. Thus, error may have occurred, in the IAR, if photos were misinterpreted by the mother.

A thorough analysis of the feasibility and acceptability, using results from the questionnaire administered at the end of the period of data collection and administrative records, will be reported separately (manuscript in preparation). These results may provide further insights into factors contributing to the wide LOA observed in this study.

Based on the results of this study, the EII (i.e. the enumerator working independently – without the aid of the mother – to interpret from the wearable camera photos foods consumed by the mother or child) did not provide a reliable estimate of DDS for mothers or children. When enumerators annotated foods consumed based on their interpretation of the photos, without the assistance of the mother, they consistently underestimated the variety of foods consumed by both mothers and children compared with the mother-assisted IAR and WFR (see online Supplementary Table S2a and b). For example, based on enumerators' review of photos alone, only 35 % of mothers consumed animal source foods compared with 62 % when the mother-assisted IAR was used (and 65 % in the WFR). For children, only 36 % (EII) compared with 60 % (IAR) and 59 % (WFR) was estimated to have consumed pulses. This suggests that study participants themselves are crucial in interpreting wearable camera images for the purposes of estimating DDS.

DDS and MDD estimated via IAR on observation days were similar to those estimated on non-observation days. The results of this study therefore indicate low reactivity to observation. However, the results achieved in this study may still reflect a higher level of agreement between the 24HR and IAR *v.* WFR than might otherwise be expected. Owing to higher percentage of breast-feeding children included in the analyses compared with those lost to the study, children in this study may have remained in closer proximity to their mother, thus enabling more consistent monitoring of the child's dietary intake. Also, mothers may have been more vigilant of the child's food intake due to reactivity to the wearable camera.

Overall, dietary quality for mothers and children in this study population was poor (see online Supplementary Table S2a and b). Consumption of most nutrient-dense foods, such as dairy products, eggs, and dark leafy greens and other vitamin A-rich fruits and vegetables by mothers and children, respectively, was low. Data collection was conducted during the dry season, and, consequently, vitamin A-rich fruit consumption may have been lower than that in other seasons. Even though the consumption of animal source foods is relatively high, due to the widespread consumption of small fish, less than half of mothers and <60 % of children achieved the MDD.

This study was conducted in Bugiri and Kamuli Districts in the Busoga Region of Eastern Uganda where 29 % of children under 5 years of age are stunted and 7 % of women are underweight[56]. In Busoga Region, 66 % of women are illiterate, only 12 % of women have completed primary school and more than a third have no regular access to radio, television or the newspaper[56]. By comparison, in our study, less than half of the participants were literate and just over one-third had only a primary school level of education. Our results show markedly better child diet

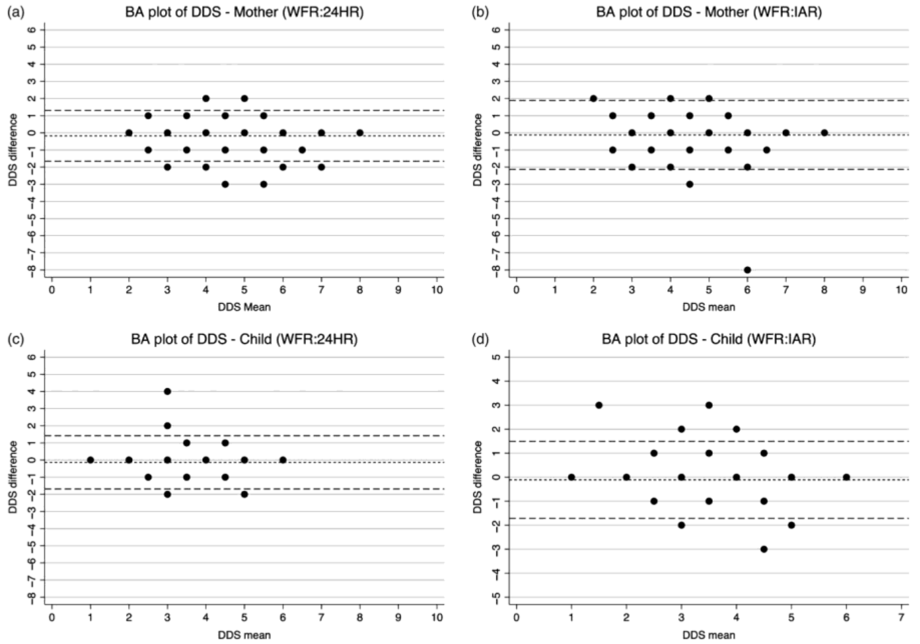


Fig. 3. Bland–Altman (BA) plots of maternal and child dietary diversity (DDS) score difference *v.* the mean. WFR, weighed food record; 24HR, 24-h recall; IAR, image-assisted recall.

Table 3. Inter-method comparisons of the relative bias and limits of agreement (LOA)
(Mean values and standard deviations)

	Mother				Child			
	Relative bias*	LOA†		LOA difference	Relative bias*	LOA†		LOA difference
24HR	−0·1801	−1·6615	1·3012	2·9627	−0·1358	−1·6883	1·4167	3·1050
IAR	−0·1227	−2·1322	1·8868	4·0190	−0·1104	−1·7130	1·4921	3·2051

24HR, 24-h recall; IAR, image-assisted recall.
* Relative mean difference.
† ± 2 SD from the relative mean difference.

Table 4. Inter-method comparisons of Cohen's κ coefficient
(κ Values and 95 % confidence intervals)

	24HR		IAR	
	Cohen's κ†	95 % CI	Cohen's κ†	95 % CI
Mother*	0·6762	0·5937, 0·7587	0·5868	0·4940, 0·6796
Child*	0·5989	0·4930, 0·7048	0·5945	0·4922, 0·6967

24HR, 24-h recall; IAR, image-assisted recall; DDS, dietary diversity score.
* Mother DDS is out of ten food groups; child DDS is out of seven food groups.
† Using weighted Cohen's κ for interrater agreement. Landis & Koch[52] suggest the following benchmark scale for interpreting the κ statistic: <0·00 poor; 0·00–0·20 slight; 0·21–0·40 fair; 0·41–0·60 moderate; 0·61–0·80 substantial; 0·81–1·00 almost perfect.

quality (MDD = 55 %) compared with that reported by the most recent Demographic and Health Surveys for the Busoga region (MDD = 31 %) (DDS was not reported)[56]. Relatively higher levels of education and literacy among mothers in this study may be a factor in higher-than-expected child MDD. Participation in the Sasakawa programme, seasonality, method of data collection or secular changes in food consumption patterns may also contribute to differences in the prevalence of children having achieved MDD.

Limitations

This study set out to pilot test and evaluate the potential of using an inexpensive LLWC to estimate the DDS and MDD of women and young children. Our hypothesis was that prospectively capturing food consumption data would reduce systematic and random errors inherent to dietary recalls and reduce respondent/interviewer burdens inherent to WFR and might allow accurate dietary diversity data collection at scale for programmatic purposes in rural low-income country (LIC) contexts. Our results indicate that, although the relative biases of both the 24HR and IAR were low, the high individual-level error observed in both methods may be expected to attenuate associations between DDS and outcomes measured. Therefore, where DDS are estimated via 24HR or IAR data, the true associations between DDS and outcomes may be stronger than they appear as a result of misreporting of food group consumption by a large proportion of the population.

In the design of the IAR protocol used in this study, several trade-offs were made. To keep equipment costs low (for LIC contexts), human interaction was required at every step of data processing. For example, although the wearable cameras are fully automated, annotation of the photos (e.g. foods consumed) required for analyses was paper-based and labour-intensive. Humans acting as the bridge between information technology systems provide ample opportunity for error and loss of data.

There were a few common scenarios in which, for the IAR, foods consumed were based solely on recall and were not, in actuality, image-assisted. For example, for reasons of enumerator safety, the wearable cameras often had to be removed before the end of the participants' day, and consequently food preparation and cooking, eating and feeding activities at the end of the day were often missed. Also, the camera was attached to the mother, so there was no visual record of foods consumed by the child under the care of someone else or when the mother was not facing her child. For this reason, we would have expected a more accurate DDS estimate via IAR for mothers than children. However, this was not observed in any of the key indicators evaluated in this study, indicating that other logistical or technical limitations were more important factors contributing to poor agreement. For example, determining ingredients – especially nutrient-dense ingredients commonly consumed in small quantities – posed a serious challenge. In this study, it was rare to see the addition of milk – for example, into tea or porridge – in photos. Milk was either added by someone other than the mother or otherwise added off camera. Milk and other ingredients may be stored in non-descript, solid-coloured containers, and mothers and children commonly drink from solid-coloured plastic mugs, making it difficult to determine the contents (e.g. to differentiate black tea from milk tea) once served. Differentiating white from yellow tubers, which have different nutrient values, was also a challenge.

In this Ugandan context, food is usually prepared at the family level and can take hours to cook, with ingredients being added long before consumption, and by various people in the household throughout the day. Also, it was a common practice for mothers in this study to prepare food, eat and feed children while seated on the ground, at an angle such that food consumption was not captured on camera. Food consumption and preparation steps may have also been missed between 30-s photo increments. Fruit, in particular, is often picked and consumed quickly, appearing in just 1 or 2 frames (out of 1800), or not at all.

Finally, where there is no electricity, pictures taken before dawn, after dark or inside the kitchen – where a lot of cooking occurs – are too dark to see, and movement of the camera can render photos indecipherable. Addressing these logistical and technical limitations may improve the relative validity of the IAR for estimating DDS and MDD for mothers and children. For children, the IAR may only be an appropriate method for assessing dietary diversity when the caregiver wearing the camera exclusively feeds the child, or for children under 12 months of age who are less mobile and require more assistance during feeding.

Conclusion

The 24HR and IAR performed similarly in estimating maternal and child DDS in this rural LIC context. For both methods and populations, there was low systematic bias. Both 24HR and IAR provided an accurate estimate of median DDS at the population level, although they both tended to overestimate the percentage of mothers and children achieving the MDD. However, importantly, this first-ever study to quantify the extent of measurement error inherent in recall methods for estimating DDS suggests that the degree of attenuation may be greater than previously recognised. Given the high LOA observed here, true associations between DDS – where estimated via 24HR or IAR – and outcomes measured may be stronger than they appear. These results, however, suggest that unless the validity of the IAR can be improved, for reasons of utility, future studies should continue to use data collected using the 24HR to estimate DDS and MDD. The time required for both data collection and processing was substantially lower for the 24HR than the IAR.

Future studies should endeavour to quantify the amount of attenuation due to misreporting food group consumption inherent to common methods for assessing DDS, and to investigate factors associated with these errors across different country contexts. As an early prototype tailored to LIC settings, the IAR performed similarly to the 24HR for estimating DDS. Further research and development to address the logistical and technical challenges identified in this study are needed to fully capitalise on the strengths of LLWC for prospectively – and passively – capturing the consumption of food/beverage in a LIC context. Additional studies are needed to determine whether active photography, where participants are instructed to photograph foods when they are consumed and the ingredients added to individual recipes, better addresses the challenges of passive photography identified in this study. Future research should also seek to exploit the unique capability of wearable cameras to simultaneously gather data related to food intake and other factors driving nutrition outcomes (e.g. time allocation, care and feeding practices, availability and accessibility to food, and cleanliness of the environment) to better understand their associations and inform the design and evaluation of nutrition-sensitive programmes in LICs.

Acknowledgements

The authors thank Stephen Young (Natural Resources Institute, UK) for his assistance with statistical analyses, Aiden Doherty (University of Oxford) and Paul Kelly (University of Edinburgh) for their advice regarding the analysis of wearable camera images, the team of enumerators and supervisors who spent tireless days and nights in the field, and the mother participants and their families who made this research possible.

This study was financially supported by an Innovative Methods and Metrics for Agriculture and Nutrition Actions (IMMANA) grant, funded by UK AID (J. N., E. L. F., K. W., grant number IMMANA 2.05). IMMANA had no role in the design, analysis or writing of this article.

A. L. S. B., L. P., J. P., L. F., G. V., J. N., K. W. and E. L. F. contributed to the study design and tool development. A. L. S. B., G. V., J. N., K. W. and E. L. F. were involved in data collection and study management. G. V., B. D. G. and N. S. entered data and contributed to interpretation of the data. A. L. S. B. carried out the validation analyses, with oversight from L. P. and E. L. F., P. K. contributed to the data analysis. A. L. S. B., L. P. and E. L. F. interpreted the results. A. L. S. B. wrote the manuscript. L. P., L. F., P. K., J. N., K. W. and E. L. F. edited the manuscript. All authors read and approved the final version of the manuscript.

The authors declare that there are no conflicts of interest.

Supplementary material

For supplementary material referred to in this article, please visit https://doi.org/10.1017/S0007114520003530

References

1. United Nations (2010) *The Millennium Development Goals Report 2010*. New York: United Nations Department of Economic and Social Affairs (DESA).
2. UNICEF (1990) *Strategy for Improved Nutrition of Infant and Women in Developing Countries*. New York: UNICEF.
3. Black RE, Alderman H, Bhutta ZA, *et al.* (2013) Maternal and child nutrition: building momentum for impact. *Lancet* **382**, 372–375.
4. United Nations Children's Fund (UNICEF), World Health Organization & International Bank for Reconstruction and Development/The World Bank (2020) *Levels and Trends in Child Malnutrition: Key Findings of the 2020 Edition of the Joint Child Malnutrition Estimates*. Geneva: UNICEF, WHO and the World Bank Group.
5. Livingstone MBE, Robson PJ & Wallace JMW (2004) Issues in dietary intake assessment of children and adolescents. *Br J Nutr* **92**, S213–S222.
6. Burrows TL, Martin RJ & Collins CE (2010) A Systematic review of the validity of dietary assessment methods in children when compared with the method of doubly labeled water. *J Am Diet Assoc* **110**, 1501–1510.
7. Subar AF, Freedman LS, Tooze JA, *et al.* (2015) Addressing current criticism regarding the value of self-report dietary data. *J Nutr* **145**, 2639–2645.
8. Foster E & Bradley J (2018) Methodological considerations and future insights for 24-hour dietary recall assessment in children. *Nutr Res* **51**, 1–11.
9. Working Group on Infant and Young Child Feeding Indicators (2006) *Developing and Validating Simple Indicators of Dietary Quality and Energy Intake of Infants and Young Children in Developing Countries: Summary of Findings from Analysis of 10 Data Sets*. Washington, DC: FANTA, FHI 360.
10. Arimond M, Wiesmann D, Becquey E, *et al.* (2010) Simple food group diversity indicators predict micronutrient adequacy of women's diets in 5 diverse, resource-poor settings. *J Nutr* **140**, 2059S–2069S.
11. Arimond M, Wiesmann D, Becquey E, *et al.* (2011) *Dietary Diversity as a Measure of the Micronutrient Adequacy of Women's Diets in Resource-Poor Areas: Summary of Results from Five Sites*. Washington, DC: FANTA-2 Bridge, FHI 360.
12. World Health Organization (2010) *Indicators for Assessing Infant and Young Child Feeding Practices: Part 2: Measurement*. Geneva: WHO.
13. FAO & FHI 360 (2016) *Minimum Dietary Diversity for Women: A Guide for Measurement*. Rome: FAO.
14. Women's Dietary Diversity Project (WDDP) Study Group (2017) Development of a dichotomous indicator for population-level assessment of dietary diversity in women of reproductive age. *Curr Dev Nutr* **1**, cdn.117.001701.
15. World Health Organization (2017) *Global Nutrition Monitoring Framework: Operational Guidance for Tracking Progress in Meeting Targets for 2025*. Geneva: World Health Organization.
16. Gibson RS (2005) *Principles of Nutritional Assessment*, 2nd ed. New York: Oxford University Press.
17. Gibson RS & Ferguson EL (2008) *An Interactive 24-Hour Recall for Assessing the Adequacy of Iron and Zinc Intakes in Developing Countries*. Washington, DC: International Life Sciences Institute.
18. Thompson FE & Subar AF (2017) Dietary assessment methodology. In *Nutrition in the Prevention and Treatment of Disease*, 4th ed., pp. 5–48 [AM Coulston, CJ Boushey, MG Ferruzzi and LM Delahanty, editors]. Cambridge, MA: Academic Press.
19. Gleason PM, Boushey CJ, Harris JE, *et al.* (2015) Publishing nutrition research: a review of multivariate techniques – part 3: data reduction methods. *J Acad Nutr Diet* **115**, 1072–1082.
20. Fosgate G (2006) Non-differential measurement error does not always bias diagnostic likelihood ratios towards the null. *Emerg Themes Epidemiol* **3**, 7.
21. Nelson M & Haraldsdóttir J (1998) Food photographs: practical guidelines I. Design and analysis of studies to validate portion size estimates. *Public Health Nutr* **1**, 219–230.
22. Marjan ZM (1995) Evaluation of food photographs for 24-hour recall method. *Malays J Nutr* **1**, 95–104.
23. Venter CS, MacIntyre UE & Vorster HH (2000) The development and testing of a food portion photograph book for use in an African population. *J Hum Nutr Diet* **13**, 205–218.

24. Turconi G, Guarcello M, Berzolari FG, *et al.* (2005) An evaluation of a colour food photography atlas as a tool for quantifying food portion size in epidemiological dietary surveys. *Eur J Clin Nutr* **59**, 923–931.
25. Robson PJ & Livingstone MBE (2000) An evaluation of food photographs as a tool for quantifying food and nutrient intakes. *Public Health Nutr* **3**, 183–192.
26. Faggiano F, Vineis P, Cravanzola D, *et al.* (1992) Validation of a method for the estimation of food portion size. *Epidemiology* **3**, 379–382.
27. Small L, Sidora-Arcoleo K, Vaughan L, *et al.* (2009) Validity and reliability of photographic diet diaries for assessing dietary intake among young children. *Infant Child Adolesc Nutr* **1**, 27–36.
28. Flax VL, Thakwalakwa C, Schnefke CH, *et al.* (2019) Validation of a digitally displayed photographic food portion-size estimation aid among women in urban and rural Malawi. *Public Health Nutr* **22**, 3140–3150.
29. Huybregts L, Roberfroid D, Lachat C, *et al.* (2008) Validity of photographs for food portion estimation in a rural West African setting. *Public Health Nutr* **11**, 581–587.
30. Harris-Fry H, Paudel P, Karn M, *et al.* (2016) Development and validation of a photographic food atlas for portion size assessment in the southern plains of Nepal. *Public Health Nutr* **19**, 2495–2507.
31. Amougou N, Cohen E, Mbala ML, *et al.* (2016) Development and validation of two food portion photograph books to assess dietary intake among adults and children in Central Africa. *Br J Nutr* **115**, 895–902.
32. De Bruyn J, Msuya J & Ferguson E (2019) Evaluating pictorial charts as a means of collecting participant-recorded data on household dietary diversity in low-literacy communities in Tanzania. *Br J Nutr* **122**, 1432–1440.
33. Thakwalakwa CM, Kuusipalo HM, Maleta KM, *et al.* (2012) The validity of a structured interactive 24-hour recall in estimating energy and nutrient intakes in 15-month-old rural Malawian children. *Matern Child Nutr* **8**, 380–389.
34. Kaczkowski CH, Jones PJH, Feng J, *et al.* (2000) Four-day multimedia diet records underestimate energy needs in middle-aged and elderly women as determined by doubly-labeled water. *J Nutr* **130**, 802–805.
35. Norman Å, Kjellenberg K, Torres Aréchiga D, *et al.* (2020) “Everyone can take photos.” Feasibility and relative validity of phone photography-based assessment of children's diets – a mixed methods study. *Nutr J* **19**, 50.
36. Rollo M, Ash S, Lyons-Wall P, *et al.* (2015) Evaluation of a mobile phone image-based dietary assessment method in adults with type 2 diabetes. *Nutrients* **7**, 4897–4910.
37. Ashman A, Collins C, Brown L, *et al.* (2017) Validation of a smartphone image-based dietary assessment method for pregnant women. *Nutrients* **9**, 73.
38. Kelly P, Marshall SJ, Badland H, *et al.* (2013) An ethical framework for automated, wearable cameras in health behavior research. *Am J Prev Med* **44**, 314–319.
39. Gemming L, Doherty A, Kelly P, *et al.* (2013) Feasibility of a SenseCam-assisted 24-h recall to reduce under-reporting of energy intake. *Eur J Clin Nutr* **67**, 1095–1099.
40. Gemming L (2015) *Image-Assisted Dietary Assessment: Evaluating the Potential of Wearable Cameras to Enhance Self-Report in the 24-Hour Dietary Recall Method*. Auckland, New Zealand: The University of Auckland.
41. Pettitt C, Liu J, Kwasnicki RM, *et al.* (2016) A pilot study to determine whether using a lightweight, wearable micro-camera improves dietary assessment accuracy and offers information on macronutrients and eating rate. *Br J Nutr* **115**, 160–167.
42. Gemming L, Rush E, Maddison R, *et al.* (2015) Wearable cameras can reduce dietary under-reporting: doubly labelled water validation of a camera-assisted 24 h recall. *Br J Nutr* **113**, 284–291.
43. Gemming L & Ni Mhurchu C (2016) Dietary under-reporting: what foods and which meals are typically under-reported? *Eur J Clin Nutr* **70**, 640–641.
44. Smith M, Stanley J, Signal L, *et al.* (2019) Children's healthy and unhealthy beverage availability, purchase and consumption: a wearable camera study. *Appetite* **133**, 240–251.
45. Kamar M, Evans C & Hugh-Jones S (2019) Factors influencing British adolescents' intake of whole grains: a pilot feasibility study using SenseCam assisted interviews. *Nutrients* **11**, 2620.
46. Martin Bland J & Altman D (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **327**, 307–310.
47. Giavarina D (2015) Understanding Bland Altman analysis. *Biochem Medica* **25**, 141–151.
48. de Vet HCW, Terwee CB, Mokkink LB, *et al.* (2011) *Measurement in Medicine: A Practical Guide. Practical guides to biostatistics and epidemiology*. Cambridge, New York: Cambridge University Press.
49. Bland JM & Altman DG (1999) Measuring agreement in method comparison studies. *Stat Methods Med Res* **8**, 135–160.
50. World Health Organization (2006) *The WHO Child Growth Standards*. Geneva: WHO.
51. Schreiner M (2016) *Uganda 2012 Poverty Probability Index (PPI)*. New Haven, CT: Innovations for Poverty Action (IPA).
52. Landis JR & Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* **33**, 159.
53. Farooq M, Doulah A, Parton J, *et al.* (2019) Validation of sensor-based food intake detection by multicamera video observation in an unconstrained environment. *Nutrients* **11**, 609.
54. 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.
55. Fuglesang A (1973) *Applied Communication in Developing Countries: Ideas and Observations*. Reprint. Uppsala: Dag Hammarskjöld Foundation.
56. Uganda Bureau of Statistics & ICF (2016) *Uganda Demographic and Health Survey 2016*. Kampala & Rockville, MD: UBOS and ICF.

SUPPLEMENTARY MATERIALS

Supplementary Table 1a. Inter-method comparisons of the median dietary diversity scores (DDS) (extended).

	WFR		24HR				EII				IAR				NON OBS EII†		NON OBS IAR‡	
	Median	25%, 75%	Median	25%, 75%	MD	P	Median	25%, 75%	MD	P	Median	25%, 75%	MD	P	Median	25%, 75%	Median	25%, 75%
Mother*	(n=163)		(n=161)				(n=159)				(n=163)				(n=133)		(n=135)	
	4	3,5	4	4,5	0	0.0017	3	2,4	1	<0.001	4	4,5	0	0.2083	3	2,4	5	4,6
Child*	(n=163)		(n=162)				(n=161)				(n=163)				(n=134)		(n=135)	
	4	3,4	4	3,4	0	0.0243	3	2,3	1	<0.001	4	3,4	0	0.0500	2	2,3	4	3,4

WFR, weighed food record; 24HR, 24-hour recall; EII, enumerator image interpretation; IAR, image-assisted recall; NON-OBS EII, non-observation day EII; NON-OBS IAR, non-observation day IAR; MD, median of the DDS differences (versus WFR).

* Mother DDS is out of 10 food groups; Child DDS is out of 7 food groups.

† Comparisons of the medians of the DDS differences versus WFR are not applicable for the non-observation day. For mothers, DDS comparisons of NON-OBS EII versus observation day EII, and NON-OBS IAR versus observation day IAR using Wilcoxon signed rank sum test are p=0.3842 and p=0.2862, respectively.

‡ Comparisons of the medians of the DDS differences versus WFR are not applicable for the non-observation day. For children, DDS comparisons of NON-OBS EII versus observation day EII, and NON-OBS IAR versus observation day IAR using Wilcoxon signed rank sum test are p=0.0267 and p=0.5243, respectively.

Supplementary Table 1b. Inter-method comparisons of the percentage achieving minimum dietary diversity (MDD) (extended).

	WFR		24HR				EII				IAR				NON OBS EII†		NON OBS IAR‡	
	n	%	n	%	MD	P	n	%	MD	*P*	n	%	MD	*P*	n	%	n	%
Mother*	(n=163)		(n=161)				(n=159)				(n=163)				(n=133)		(n=135)	
	67	41.1	76	47.2	-0.0621	0.0412	19	12.0	0.2830	<0.001	69	42.3	-0.0123	0.6698	20	15.0	73	54.1
Child*	(n=163)		(n=162)				(n=161)				(n=163)				(n=134)		(n=135)	
	90	55.2	97	59.9	-0.0494	0.1701	24	14.9	0.4037	<0.001	95	58.3	-0.0307	0.3692	18	13.4	76	56.3

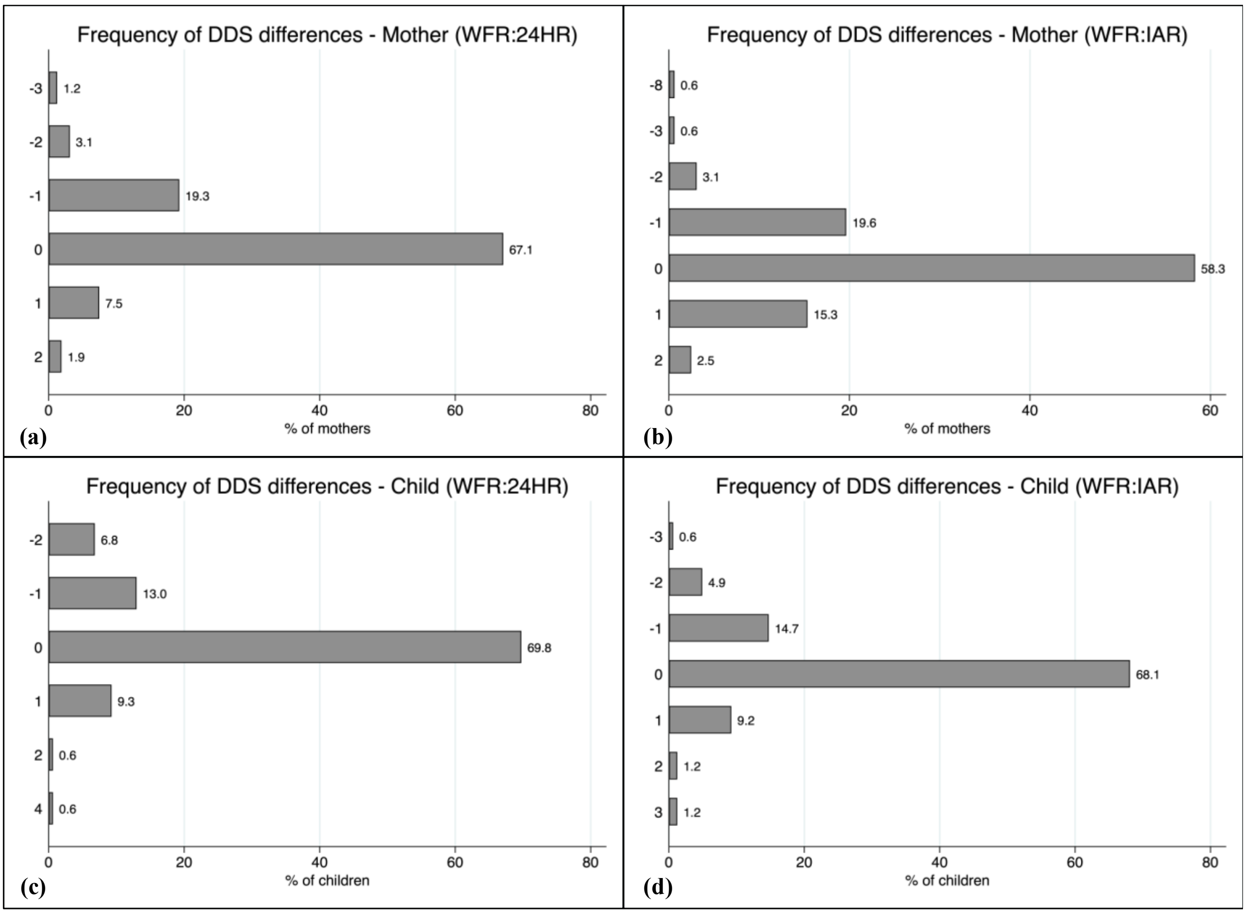
WFR, weighed food record; 24HR, 24-hour recall; EII, enumerator image interpretation; IAR, image-assisted recall; NON-OBS EII, non-observation day EII; NON-OBS IAR, non-observation day IAR; MD, mean MDD proportion difference (versus WFR).

* Mother MDD is DDS >= 5; Child MDD is DDS >= 4.

† Comparisons of the mean MDD proportion difference versus WFR are not applicable for the non-observation day. For mothers, MDD comparisons of NON-OBS EII versus observation day EII, and NON-OBS IAR versus observation day IAR using McNemar's test are p=0.5637 and p=0.1161, respectively.

‡ Comparisons of the mean MDD proportion difference versus WFR are not applicable for the non-observation day. For children, MDD comparisons NON-OBS EII versus observation day EII, and NON-OBS IAR versus observation day IAR using McNemar's test are p=0.4328 and p=0.3428, respectively.

Supplementary Figure S1. Frequency of DDS differences



Supp. Fig. 1. Frequency of maternal and child diet diversity score (DDS) differences via 24-hour recall (24HR) or image-assisted recall (IAR) versus weighed food record (WFR). (a) Frequency of DDS differences for the mother via 24HR vs. WFR, (b) Frequency of DDS differences for the mother via IAR vs. WFR, (c) Frequency of DDS differences for the child via 24HR vs. WFR, and (d) Frequency of DDS differences for the child via IAR vs. WFR.

Supplementary Table 2a. Inter-method comparisons of the food groups consumed by the mother.

	WFR		24HR				EII				IAR			
	n	%	n	%	MD	P	n	%	MD	P	n	%	MD	P
Grains, white roots and tubers, and plantains	163	100.0	162	99.4	0.0061	0.3173	155	95.7	0.0432	0.0082	159	97.6	0.0245	0.0455
Pulses	66	40.5	67	41.1	-0.0061	0.7815	49	30.3	0.1049	0.0031	65	39.9	0.0061	0.7963
Nuts and seeds	53	32.5	55	33.7	-0.0123	0.4142	31	19.1	0.1296	0.0001	56	34.4	-0.0184	0.4054
Dairy	28	17.2	34	20.9	-0.0368	0.0578	11	6.8	0.1049	0.0002	36	22.1	-0.0491	0.0209
Meat, poultry and fish	106	65.0	105	64.4	0.0061	0.7055	56	34.6	0.3025	<0.001	100	61.4	0.0368	0.1088
Eggs	5	3.1	5	3.1	0.0000	1.0000	0	0.0	0.0252	0.0455	8	4.9	-0.0184	0.0833
Dark green leafy vegetables	13	8.0	14	8.6	-0.0061	0.3173	10	6.2	0.0185	0.3173	15	9.2	-0.0123	0.1573
Other vitamin A-rich fruits and vegetables	21	12.9	31	19.0	-0.0613	0.0124	17	10.5	0.0247	0.4142	34	20.9	-0.0798	0.0046
Other vegetables	155	95.1	157	96.3	-0.0123	0.3173	74	45.7	0.4938	<0.001	147	90.2	0.0491	0.0325
Other fruits	91	55.8	100	61.7	-0.0556	0.0290	85	52.8	0.0248	0.5050	101	62.0	-0.0613	0.0330

WFR, weighed food record; 24HR, 24-hour recall; EII, enumerator image interpretation; IAR, image-assisted recall; MD, mean proportion difference (versus WFR).

Supplementary Table 2b. Inter-method comparisons of the food groups consumed by the child.

	WFR		24HR				EII				IAR			
	N	%	N	%	MD	P	N	%	MD	P	N	%	MD	P
Grains, white roots and tubers, and plantains	163	100.0	160	98.2	0.0184	0.0833	149	92.0	0.0802	0.0003	161	98.8	0.0123	0.1573
Pulses	96	58.9	99	60.7	-0.0184	0.4913	59	36.4	0.2284	<0.001	98	60.1	-0.0123	0.6374
Dairy	38	23.3	42	25.8	-0.0245	0.2059	14	8.6	0.1481	<0.001	43	26.4	-0.0307	0.0956
Meat, poultry and fish	97	59.5	100	61.4	-0.0184	0.3657	52	32.1	0.2778	<0.001	98	60.1	-0.0061	0.8185
Eggs	6	3.7	6	3.7	0.0000	1.0000	1	0.6	0.0248	0.0455	9	5.5	-0.0184	0.0833
Other vitamin A-rich fruits and vegetables	32	19.6	46	28.2	-0.0859	0.0010	18	11.1	0.0864	0.0060	43	26.4	-0.0675	0.0218
Other fruits or vegetables	157	96.3	158	96.9	-0.0061	0.6547	112	69.1	0.2716	<0.001	155	95.1	0.0123	0.4795

WFR, weighed food record; 24HR, 24-hour recall; EII, enumerator image interpretation; IAR, image-assisted recall; MD, mean proportion difference (versus WFR).

ADDITIONAL MATERIALS (NOT INCLUDED IN THE MANUSCRIPT)

Additional Table 1. Proportion of mothers reported in the 24HR to have consumed food groups before and after confirmation.

	24HR before confirmation	24HR after confirmation	
	n (%)	n (%)	P*
Grains, white roots and tubers, and plantains	189 (99.5)	189 (99.5)	1.000
Pulses	79 (41.6)	81 (42.6)	0.500
Nuts and seeds	65 (34.2)	65 (34.2)	1.000
Dairy	37 (19.5)	38 (20.0)	1.000
Meat, poultry, and fish	114 (60.0)	115 (60.5)	1.000
Eggs	5 (2.7)	6 (3.2)	1.000
Dark green leafy vegetables	13 (6.9)	14 (7.4)	1.000
Other vitamin A-rich fruits and vegetables	32 (16.8)	33 (17.4)	1.000
Other vegetables	177 (93.2)	183 (96.3)	0.031
Other fruits	112 (59.3)	118 (62.4)	0.070

24HR, 24-hour recall.
*P-value of McNemar's test of the proportion for food group consumption reported in the 24HR after confirmation compared to before confirmation <0.05

In the 24HR instrument and protocol, both the mother's initial report for each group and her final report, after having been asked to confirm the foods / beverages consumed and not consumed (as described above) were entered, so that changes between the initial report and confirmation could be assessed. Mothers reported, for themselves, higher consumption of other vegetables in the 24HR after confirmation. For all other food groups, there were no significant differences before and after confirmation.

Additional Table 2. Proportion of children reported in the 24HR to have consumed food groups before and after confirmation.

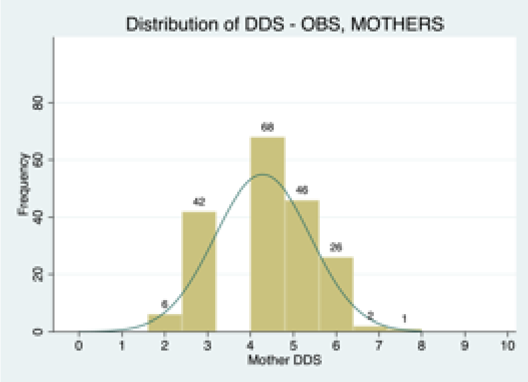
	24HR before confirmation	24HR after confirmation	
	n (%)	n (%)	P*
Grains, roots, and tubers	189 (99.5)	187 (98.4)	0.500
Legumes and nuts	114 (60.0)	118 (62.1)	0.125
Dairy products	47 (24.7)	48 (25.3)	1.000
Flesh foods	100 (52.6)	110 (57.9)	0.002
Eggs	5 (2.7)	7 (3.7)	0.500
Vitamin A-rich fruits and vegetables	48 (25.3)	51 (26.8)	0.508
Other fruits and vegetables	183 (96.3)	183 (96.3)	1.000

24HR, 24-hour recall.
*P-value of McNemar's test of the proportion for food group consumption reported in the 24HR after confirmation compared to before confirmation <0.05

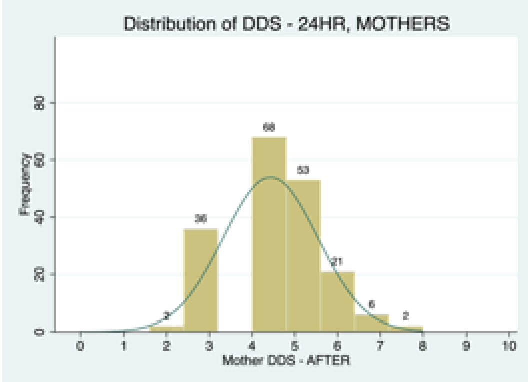
In the 24HR instrument and protocol, both the mother's initial report for each group and her final report, after having been asked to confirm the foods / beverages consumed and not consumed (as described above) were entered, so that changes between the initial report and confirmation could be assessed. Mothers reported, for their children, higher consumption of flesh foods in the 24HR after confirmation. For all other food groups, there were no significant differences before and after confirmation.

Additional Figure 1. Inter-method comparison of the distribution of DDS for mothers.

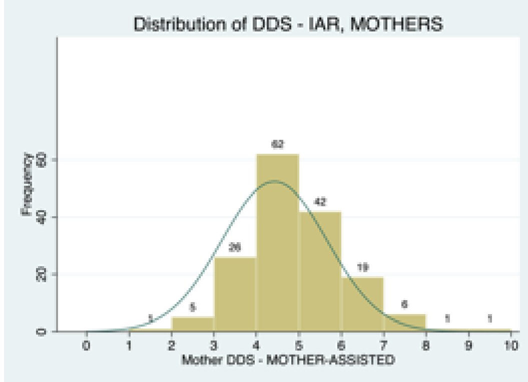
WFR



24HR



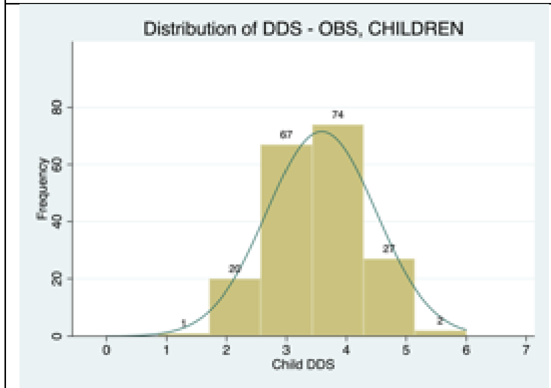
AWC-IAR



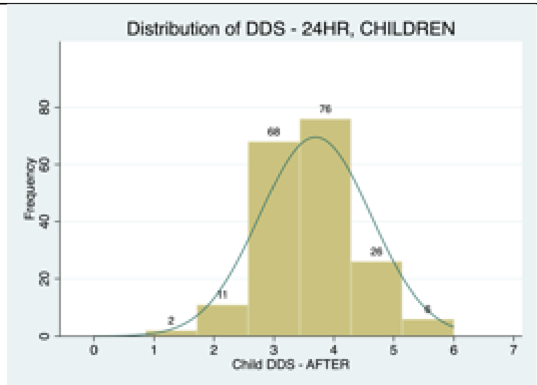
WFR, weighed food record; 24HR, 24-hour recall; AWC-IAR, automated wearable camera-based image-assisted recall.

Additional Figure 2. Inter-method comparison of the distribution of DDS for children

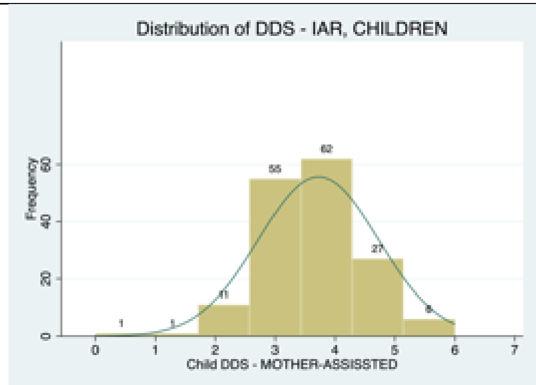
WFR



24HR



AWC-IAR



Chapter 6: Validation of an Automated Wearable Camera-Based Image-Assisted Recall Method and the 24-Hour Recall Method for Assessing Women's Time Allocation in a Nutritionally Vulnerable Population: The Case of Rural Uganda

ABSTRACT

Accurate data are essential for investigating relationships between maternal time-use patterns and nutritional outcomes. The 24 h recall (24HR) has traditionally been used to collect time-use data, however, automated wearable cameras (AWCs) with an image-assisted recall (IAR) may reduce recall bias. This study aimed to evaluate their concurrent criterion validity for assessing women's time use in rural Eastern Ugandan. Women's (n = 211) time allocations estimated via the AWC-IAR and 24HR methods were compared with OBS (criterion method) using the Bland–Altman limits of agreement (LOA) method of analysis and Cronbach's coefficient alpha (time allocation) or Cohen's κ (concurrent activities). Systematic bias varied from 1 min (domestic chores) to 226 min (caregiving) for 24HR and 1 min (own production) to 109 min (socializing) for AWC-IAR. The LOAs were within 2 h for employment, own production, and self-care for 24HR and AWC-IAR but exceeded 11 h (24HR) and 9 h (AWC-IAR) for caregiving and socializing. The LOAs were within four concurrent activities for 24HR (−1.1 to 3.7) and AWC-IAR (−3.2 to 3.2). Cronbach's alpha for time allocation ranged from 0.1728 (socializing) to 0.8056 (own production) for 24HR and 0.2270 (socializing) to 0.7938 (own production) for AWC-IAR. For assessing women's time allocations at the population level, the 24HR and AWC-IAR methods are accurate and reliable for employment, own production, and domestic chores but poor for caregiving and socializing. The results of this study suggest the need to revisit previously published research investigating the associations between women's time allocations and nutrition outcomes.

Research Paper Cover Sheet

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1407282	**Title**	Ms.
First Name(s)	Andrea		
Surname/Family Name	Spray		
Thesis Title	Validation of an automated wearable camera-based image-assisted recall method and the 24-hour recall method for assessing maternal and child dietary diversity and women's time use in rural Uganda		
Primary Supervisor	Elaine Ferguson		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	Nutrients		
When was the work published?	April 2022		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	A. L. S. B. contributed to the study design and tool development, coordinated data collection and contributed to study management, developed data entry tools, entered data and oversaw data entry research assistants, carried out the validation analyses, interpreted the results, and wrote and edited the manuscript.

SECTION E

Student Signature	[redacted]
Date	27/6/2022

Supervisor Signature	[redacted]
Date	15/6/2022

MANUSCRIPT

nutrients

MDPI

Article

Validation of an Automated Wearable Camera-Based Image-Assisted Recall Method and the 24-Hour Recall Method for Assessing Women's Time Allocation in a Nutritionally Vulnerable Population: The Case of Rural Uganda

Andrea L. S. Bulungu [1,*], Luigi Palla [2,3,4,*], Jan Priebe [5], Lora Forsythe [5], Pamela Katic [5], Gwen Varley [5], Bernice D. Galinda [1], Nakimuli Sarah [1], Joweria Nambooze [6,7], Kate Wellard [5] and Elaine L. Ferguson [1]

[1] Department of Population Health, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK; bernicegalinda9@gmail.com (B.D.M.); kigozi1995@gmail.com (N.S.); elaine.ferguson@lshtm.ac.uk (E.L.L.)
[2] Department of Public Health and Infectious Diseases, University of Roma La Sapienza, Rome 00185, Italy; luigi.palla@uniroma1.it
[3] Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK
[4] School of Tropical Medicine and Global Health, University of Nagasaki, Nagasaki 852-8102, Japan
[5] Natural Resources Institute (NRI), University of Greenwich, Chatham Maritime, Kent ME4 4TB, UK; jan.priebe@gmail.com (J.P.); l.forsythe@greenwich.ac.uk (L.F.); p.g.katic@greenwich.ac.uk (P.K.); gwendolyn.varley@slu.se (G.V.); k.wellard@greenwich.ac.uk (K.W.)
[6] Africa Innovations Institute (AfrII), Kampala P.O. Box 34981, Uganda; jngalabuzi@yahoo.com
[7] Department of Nutritional Sciences and Dietetics, Kyambogo University, Kyambogo, Kampala P.O. Box 1, Uganda; jngalabuzi@yahoo.com
* Correspondence: andrealspray@gmail.com (A.L.S.B.); luigi.palla@uniroma1.it (L.P.)

Citation: Bulungu, A.L.S.; Palla, L.; Priebe, J.; Forsythe, L.; Katic, P.; Varley, G.; Galinda, B.D.; Sarah, N.; Nambooze, J.; Wellard, K.; Ferguson, E.L. Validation of an Automated Wearable Camera-Based Image-Assisted Recall Method and the 24-Hour Recall Method for Assessing Women's Time Allocation in a Nutritionally Vulnerable Population: The Case of Rural Uganda. *Nutrients* **2022**, *14*, 1833. https://doi.org/10.3390/nu14091833

Academic Editor: Bo Chen

Received: 31 March 2022
Accepted: 26 April 2022
Published: 27 April 2022

Abstract: Accurate data are essential for investigating relationships between maternal time-use patterns and nutritional outcomes. The 24 h recall (24HR) has traditionally been used to collect time-use data, however, automated wearable cameras (AWCs) with an image-assisted recall (IAR) may reduce recall bias. This study aimed to evaluate their concurrent criterion validity for assessing women's time use in rural Eastern Ugandan. Women's (n = 211) time allocations estimated via the AWC-IAR and 24HR methods were compared with direct observation (criterion method) using the Bland–Altman limits of agreement (LOA) method of analysis and Cronbach's coefficient alpha (time allocation) or Cohen's κ (concurrent activities). Systematic bias varied from 1 min (domestic chores) to 226 min (caregiving) for 24HR and 1 min (own production) to 109 min (socializing) for AWC-IAR. The LOAs were within 2 h for employment, own production, and self-care for 24HR and AWC-IAR but exceeded 11 h (24HR) and 9 h (AWC-IAR) for caregiving and socializing. The LOAs were within four concurrent activities for 24HR (−1.1 to 3.7) and AWC-IAR (−3.2 to 3.2). Cronbach's alpha for time allocation ranged from 0.1728 (socializing) to 0.8056 (own production) for 24HR and 0.2270 (socializing) to 0.7938 (own production) for AWC-IAR. For assessing women's time allocations at the population level, the 24HR and AWC-IAR methods are accurate and reliable for employment, own production, and domestic chores but poor for caregiving and socializing. The results of this study suggest the need to revisit previously published research investigating the associations between women's time allocations and nutrition outcomes.

Keywords: validation studies; time use; methodology; wearable camera; measurement error; care practices; maternal time

1. Introduction

Time is an essential resource for caregiving, including care for women, child feeding and psychosocial and cognitive stimulation, hygiene practices, home health practices, and food preparation and storage [1]. In poor households of low-income countries, the provision of essential needs (food, water, care) relies primarily on the time and labor of household members themselves. Longstanding evidence from time-use studies around the world indicates that, despite high rates of participation in productive work, the burden of care for children and other household members remains primarily with women [2–6]. Understanding the simultaneous demands on women's time for basic survival, and the trade-offs made between time allocated to food production, food preparation, income-generating activities, home maintenance, and care of children and other members of the household, is essential for understanding the factors influencing nutrition in low-income country contexts [7–13].

Popkin (1980) first demonstrated that a mother's time spent on childcare was positively associated with child nutrition status [14] and, subsequently, the importance of time for care has been well understood to be a key factor for maternal and child nutrition and overall well-being [1]. Yet, decades of empirical studies have shown the relationship between women's time allocation and maternal and/or child nutrition to be complex. Gryoboski (1996) found that time allocated to childcare by aunts, sisters, and grandmothers was associated with increased caloric intake for children [15], whereas the association between the mother's time allocated to childcare and caloric intake was negative, and Komatsu et al. (2018) found that the association between women's time allocated to "reproductive work" and child nutrition depended on the economic status of the household [8]. Other studies have taken the opposite approach, exploring instead the influence of women's time allocated to productive work on their and/or their child's nutritional status. The results of these studies have also been conflicting. Some have found that a mother's time allocated to (agriculture) work was negatively associated with child nutritional status [16,17], whereas others have found there to be no relationship between a mother's time allocated to work and child nutritional status [12,18].

Previous studies suggest that the amount of time women allocate to both productive and reproductive work is severely undercounted [19–25], which limits our ability to accurately assess how women spend their time and its influence on women's and children's nutrition, health, and well-being [16,26–33]. The boundaries of "reproductive work" are not rigid. Caregiving responsibilities and other domestic chores tend to overlap with time allocated to income-generating activities or leisure. A large proportion of reproductive work, such as childcare, is performed while simultaneously performing other tasks [27,34,35]. Rather than time spent in "productive work" displacing time spent in "reproductive work", women tend to manage their dual roles by simply working more hours [36] and/or multitasking [37–39]. This overlap in productive and reproductive work is differentially detrimental to measuring women's time in their dual role as caregivers and income generators. Current methods for measuring time allocation have limitations that undermine efforts to accurately estimate simultaneous activities, especially in rural, low-income country contexts [27,34,40–45].

For over a century, surveys in high-income countries have typically utilized time-use diary methods (i.e., time budgets—either prospective or retrospective) and these are still considered to be reliable methods of assessing time use [46–48]. However, for low-income country contexts where literacy is low, or the sense of time does not align with western constructs, time-use diaries are infeasible [49]. In these contexts, direct observation is generally considered to be the "gold standard" for collecting time allocation data [6,22,50–56]. For time use, "direct observation" entails a researcher being at participants' homes, watching and recording everything they do. It is resource intensive, requires specialized skill, is burdensome for the participant, and is prone to various forms of reactivity bias, including the Hawthorne effect, observer-expectancy effect, and observer bias [30,57–59]. Therefore, for time-use surveys in low-income countries, the 24 h recall (24HR) method is commonly used [40,51,55,60–62]. The 24HR method is a structured interview where the interviewer asks the respondent, for time use, about all the activities they performed the previous day.

Recall methods, such as the retrospective time-use diary, the 24 h "free" recall technique utilized in the time-use module of the Women's Empowerment in Agriculture Index (WEAI), or the more conventional "stylized questionnaires", levy a heavy cognitive burden on the participant and are prone to error and bias [63–65]. Their accuracy relies upon the respondent's memory and motivation, as well as on the skill and persistence of the interviewer. Misreporting of activities may occur, either unintentionally or intentionally, due to biases (i.e., recall, interviewer, social desirability, or approval biases), low literacy or numeracy, or cognitive phenomena unique to the recall of time allocation, such as "telescoping" and the burden of aggregating time across hours of the day, days of the week, or seasons of the year [23,66–68]. In particular, recall methods poorly capture reproductive activities that are typically performed by women in rural low-income country contexts, such as: childcare and feeding, food preparation, domestic chores (fetching fuel, washing clothes, etc.), healthcare seeking, and socializing [50,69].

Historically, informal work (e.g., selling food stuffs in unregulated markets), such as is commonly performed by women in rural low-income countries, was unrecognized in large-scale labor force and time-use studies [19,21,27]. Efforts to remediate the under-reporting of informal and reproductive work have culminated in the launch, in 2019, of the International Classification of Activities for Time-Use Statistics (ICATUS-2016) [70]. These recent improvements, however, do not address the methodological limitations of traditional retrospective time allocation assessment methods for measuring work predominantly done by women in rural low-income country settings.

Automated wearable cameras (AWCs) are inexpensive technologies that prospectively and unobtrusively record activities as they are performed. AWCs have been investigated as a research method in high-income countries for assessing diet [71–83], physical activity [84–89], and the food environment [90–95]. Only limited research with AWCs has been undertaken in low- or middle-income country contexts. They have been evaluated for assessing diet in Tonga [96], China [97], and Uganda [98], and childcare practices in South Africa and Nepal [99].

An IAR is a method using photographs, either automatically generated from an AWC or taken by the participant, as an autobiographical memory cue (recall trigger) to help respondents reconstruct key details from their previous day [100–103]. Only a few studies using an AWC-based IAR to assess time allocation have been conducted, all in the UK [102,104,105]. Bulungu et al. (2021) identified several challenges unique to rural low-income country settings which may affect their performance in assessing women's time use, including a subject's difficulty in interpreting the wearable camera's first-person perspective photos, activities happening outside the camera's field of vision, and poor lighting [98].

This study was, therefore, undertaken in rural Eastern Uganda to evaluate the concurrent criterion validity, for assessing women's time use, comparing both an AWC-based image-assisted recall (IAR) method and the 24HR method to direct observation (criterion method). No study, to our knowledge, has examined the criterion validity, for assessing women's time use in a low-income country context, of the 24HR method or an AWC-based method in either free-living or controlled settings. In addition, as described by Bulungu et al. (2021), in this study population the median dietary diversity score was 4 for both women (out of 10 food groups) and children (out of seven food groups), with only 41% and 42% of women and children achieving the minimum dietary diversity, respectively, which indicates this population represents a nutritionally vulnerable one.

2. Materials and Methods

2.1. Study Design

This study was nested within a cross-sectional study of women with a child aged between 12 and 23 months inclusive (n = 211), to examine the impact of a labor-saving technology on women's time for childcare, food preparation, and dietary practices. The study was conducted between January and February 2018 in Bugiri and Kamuli Districts, Eastern Region, Uganda.

In our study, women's time allocation was assessed, for the same day, using three concurrent methods: (1) direct continuous observation (15 h), (2) 24HR, and (3) IAR using photos captured via an AWC. An IAR is a method using photographs, either automatically generated from a wearable camera or taken by the participant, as an autobiographical memory cue (recall trigger) to help respondents reconstruct key details from their previous day [100–103].

Data were collected over five consecutive days, following one of two possible patterns, as presented in Supplementary Table S1. Specifically, for both patterns, on day 1, eligibility was confirmed, a structured questionnaire was administered, and anthropometric data were collected for all participants. For half of the study participants, on day 2, time allocation data were collected using direct observation and recorded on the AWC attached to the respondent. On day 3, a 24HR was administered, followed by an IAR using photos captured on day 2 by the AWC. On day 4, time allocation data were again recorded via AWC only (i.e., no observation). On day 5, an IAR was administered using photos captured on day 4 by the AWC. The other half of the study participants began with the AWC only (i.e., days 2 and 4 were switched) and ended with all three methods (i.e., days 3 and 5 were switched). For all participants, on the 5th day, a final structured questionnaire was also administered. Time allocation data collection was distributed across all days of the week at the population level to account for a day-of-the-week effect, and for each respondent, the enumerator assigned to conduct the direct observation was different from the enumerator assigned to administer the 24HR and IAR.

Ethical approval was obtained from the [location masked for blind review] (A24ES), [location masked for blind review] Research Ethics Committee (Project ID: 1420), and [location masked for blind review] Ethics Committee (Project ID: B0501). Following community sensitization, verbal explanation of the study, and demonstration of the AWC, written consent (signature or thumb print) was obtained from all respondents who participated in our study.

2.2. Participants and Sampling

Twenty-two villages were purposefully selected, for this study, of which eleven had access to labor-saving technology and eleven did not. These villages participated in the Sasakawa Global 2000 Uganda (SG2000 Uganda) country program (the local implementing partner for the parent study). The sample size calculation (n = 264; 22 communities, 12 households per community) was based on requirements of the main study within which this current study was nested. This sample size was deemed sufficient for the current validation study, using the Bland–Altman (BA) method of analyses [106–108].

The sampling frame, for each village, was a household listing of all women with a child born between 1 January 2017 and 1 May 2017 (children aged 12 to 23 months at the time of data collection). These lists were generated by the SG2000 community-based facilitators. Twelve mother–child dyads in each village were randomly selected to participate in the study. Substitutions were made, as needed, until 12 mother–child dyads who met the inclusion/exclusion criteria were recruited. Mother–child dyads were excluded if the child was less than 12 months or greater than 23 months of age, was not yet eating solid foods on a regular basis, or was a multiple-birth child; the mother was unable to communicate in Lusoga, Luganda, or English; either the mother or child had a severe disability; the mother was not the biological mother of the child; the mother was a co-wife with a selected mother; or either the mother or child was not available for the duration of the study.

2.3. Instruments and Protocol

The enumerators administered two structured questionnaires to the respondent. The first questionnaire collected information on household socio-demographics and assets, and factors related to women's empowerment. The second questionnaire collected information on household mobile phone access and ownership, and perceptions of their experiences with each of the three time allocation data collection methods.

For the criterion time allocation assessment method (i.e., direct observation), enumerators recorded all activities undertaken by the respondent in 15 min intervals ("timeslots") from approximately 06:00 to 21:00, using a structured instrument comprising 44 activities.

On the day after the observation day, a multiple-pass 24HR was administered to the respondent to collect information on all activities undertaken by the respondent on the previous day. In the first pass, the respondent was asked to list everything she did the previous day; in the second pass, additional details about each activity and any concurrent activities were recorded. The time and duration of each activity were recorded in 15 min increments. In the third pass, the enumerator confirmed with the respondent that each activity was recorded accurately. The 24HR protocol was based on a module developed for the WEAI, which was itself based on the Lesotho Time Budget Study [65,109].

On the observation day, a small, lightweight AWC (iON SnapCam Lite, dimensions 42 × 42 × 13 mm) was attached to a t-shirt worn by the respondent at approximately 06:00 and removed at approximately 21:00. Participants were instructed to wear the AWC while continuing their usual activities, covering or removing the camera as needed for privacy. The AWC automatically recorded a picture every 30 s, storing all photos (approximately 1800) on a memory card.

The following day, an enumerator first reviewed the photos captured by the AWC on a tablet and annotated the activities she thought—based on the photos—were undertaken by the respondent, i.e., the enumerator image interpretation (EII). Based on her interpretation of the photos, the enumerator demarcated the series of activities for review later that day with the respondent. Upon meeting with the respondent, the enumerator first administered the 24HR. The enumerator then administered the IAR by first reviewing the AWC photos with the respondent on the tablet (16GB Samsung Galaxy Tab 3 with a 10" screen, using Simple Gallery software for image display). During this interview, the enumerator used "verbal probing" to elicit from the participant additional relevant information about the activities performed, for example, to elaborate on what she was doing, who she was with, where she was going, and why [110,111]. The enumerator revised her original annotations (i.e., the EII) of activities undertaken by the respondent, as needed, based on the respondent's feedback.

The IAR protocol was adapted from one described by Kelly et al. (2015). The protocol followed ethical guidelines for AWC research to ensure privacy of the participants was maintained [112]. All protocols were pilot tested and refined prior to the start of the study.

2.4. Data Processing

The number of minutes allocated to each of 44 activities recorded over the fifteen-hour period was calculated for each respondent and for each of the 3 data collection methods, in 15 min intervals ("timeslots"). The discrete activities were categorized into the nine mutually exclusive ICATUS-2016 major divisions ("activity groups"): (1) employment and related activities ("employment"), (2) production of goods for own final use ("own production"), (3) unpaid domestic services for household and family members ("domestic chores"), (4) unpaid caregiving services for household and family members ("caregiving"), (5) unpaid volunteer, trainee, and other unpaid work, (6) learning, (7) socializing and communication, community participation, and religious practice ("socializing"), (8) culture, leisure, mass media, and sports practices ("leisure"), and (9) self-care and maintenance ("self-care"), as presented in Supplementary Table S2 [70]. When individuals were observed performing more than one activity concurrently, the activities were given equal weight such that no activity was deemed "primary" or "secondary".

Of the 44 activities tracked, four were considered to be "simultaneous", i.e., they could be performed while also performing other activities: care of the index child, care of other children or adults, chatting with friends or family, and watching TV or listening to the radio. When just one activity was performed in a timeslot, the activity performed counted for the entire 15 min. The simultaneous activities were always credited the full 15 min. However, for all other activities, when more than one activity was performed per timeslot, the 15 min were evenly distributed across the activities performed. For example, if in a 15 min timeslot, the participant was snacking (self-care) and then started preparing food (domestic chores) while feeding the index child (caregiving), caregiving—a simultaneous activity—was credited 15 min and self-care and domestic chores were each credited 7.5 min.

The proportion of the study population living below USD 1.25/day was calculated using the Uganda 2012 Poverty Probability Index (PPI) with data collected via the respondents' questionnaires [113].

2.5. Data Analysis

The primary outcome variables analyzed were the total minutes allocated to each of the nine ICATUS-2016 activity groups and the median number of concurrent activities performed across all 15 min timeslots. Data were analyzed using Stata/SE version 15.1. *P*-values less than 0.05 were considered significant for all tests. Cases with incomplete data for any of the three methods (observation, 24HR, or IAR) were eliminated from analysis. Key socio-demographic characteristics for participating and missing households were compared using the Mann–Whitney U two-sample statistic for continuous data, and the Fisher exact test for categorical data.

Due to inter-participant differences in actual observation start and end times and technical challenges with insufficient light in the early morning and evening, the analyses were limited to the 12 h period from 8 a.m. to 8 p.m. to retain as many cases as possible with complete data. The Wilcoxon signed rank sum test was used to compare the distributions of time allocation obtained via the criterion method (observation) versus the 24HR, IAR, or EII. The median time allocated for only those women partaking in each activity was also calculated and compared using the Wilcoxon signed rank sum test. Women's time allocations estimated via EII and IAR were also calculated for the non-observation day and compared to the corresponding estimates for the observation day using the Wilcoxon signed rank sum test. The Wilcoxon signed rank sum test was also used to compare the distributions of the median number of concurrent activities obtained via the criterion method (observation) versus the 24HR or IAR.

The inter-tool agreement between the criterion method (observation) and 24HR or IAR was assessed using the Bland–Altman limits of agreement (LOA) method for each ICATUS-2016 major division (minutes/d) [106]. Specifically, for each individual, the differences between the methods (the criterion measure of time allocation minus the time allocation estimated using either 24HR or IAR) versus the mean of the two methods were plotted; the bias and the 95% LOA (mean difference ± 2 SD of the differences) were estimated. The numbers of participants for whom the differences between the two methods were greater or less than zero were also calculated. The Bland–Altman LOA approach was used to assess inter-method agreement for estimating the median number of concurrent activities.

Time allocations estimated via the 24HR and IAR methods against the reference method were compared using Cronbach's (reliability) coefficient alpha. Cronbach's coefficient alpha was interpreted as follows: <0.70 unacceptable; >0.70 acceptable; >0.80 moderate; 0.90–0.95 high; >0.95 suspect [114]. The inter-method reliability (24HR and IAR methods against the criterion method) for estimating the median number of concurrent activities was compared using the weighted Cohen's κ coefficient. Cohen's κ coefficient was interpreted as follows: <0·00 poor agreement; 0·00–0·20 slight agreement; 0·21–0·40 fair agreement; 0·41–0·60 moderate agreement; 0·61–0·80 substantial agreement; 0·81–1·00 almost perfect agreement [115,116].

3. Results

3.1. Characteristics of the Sample

Overall, 211 women were recruited into the study. Among those recruited, six participants voluntarily withdrew, and 30 participants were eliminated from analysis due to incomplete data (Figure 1). Characteristics of the study population are presented and compared with participants who were lost to follow-up or excluded from the analyses in Table 1. These comparisons show some differences between them, including child breastfeeding status (60% for participants vs. 41% for non-participants) and alternative childcare provided exclusively by persons aged thirteen years or older (39% for participants vs. 64% for non-participants). The median household size was six members, and nearly one quarter of participating households lived below USD 1.25/day. Most participating respondents were married and between the ages of 20 and 29 years. Nearly two-thirds of participating respondents had not completed primary school, and just under one half were literate. Most respondents were pregnant, breastfeeding, or both.

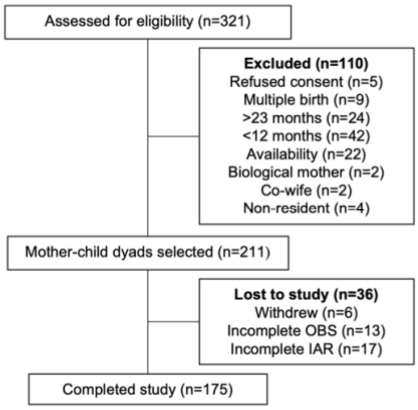


Figure 1. Study population. OBS, observation; IAR, image-assisted recall.

Table 1. Characteristics of households, mothers, and children participating in and excluded from the analysis.

	Participating				Participants Excluded from Analyses				P
	n	%	Median	25th, 75th	n	%	Median	25th, 75th	
Households									
Number of household members			6	5, 8			5	4, 7	0.1017
Living below USD 1.25/day (2005 PPP)	140	24.1			23	21.2			0.086
Mothers									
Age (years)									0.0449
15–19	18	10.3			6	16.7			0.521
20–29	105	60	26	22, 30	23	63.9	23	20, 28	
30–39	44	25.1			6	16.7			
40–49	8	4.6			1	2.8			
Marital status									
Single	19	10.9			2	6.1			0.833
Married or co-habiting	147	84.5			30	90.9			
Level of education									
None or primary incomplete	106	60.1			18	54.6			0.378
Primary complete	62	35.4			12	36.4			
Secondary complete	5	2.9			2	6.1			
Can read and write	82	48			20	60.6			0.127
Maternity status									
Pregnant	25	14.9			8	23.5			0.16
Breastfeeding	110	62.9			17	47.2			0.061
Pregnant or breastfeeding	129	73.7			23	63.9			0.16
Children									
Age (months)									0.9001
12–17	104	59.8	16.7	14.8, 20.0	19	54.3	17.7	14.8, 19.6	0.338
18–23	70	40.2			16	45.7			
Sex									
Female	78	44.6			20	57.1			0.12
Male	97	55.4			15	42.9			
Ever breastfed	172	99.4			31	96.9			0.288
Currently breastfeeding	103	59.5			13	40.6			0.037
Child caregivers									0.2597
No alternative caregivers	16	9.1	3	2, 4	3	8.3	3	2, 4	0.588
All child caregivers > 13 years	68	38.9			23	63.9			0.005

PPP, purchasing power parity; P, *p*-value using Mann–Whitney U test to compare the medians and Fisher's exact test to compare the categorical data.

The median age of participating children was 17 months, and there were slightly more males (55%) than females (45%). Nearly all children were initially breastfed, although just 60% were breastfeeding at the time of data collection. Among the study participants, over 90% of children were cared for by at least one alternative caregiver (other than their mother), of which more than 60% included at least one alternative caregiver who was less than 13 years of age.

3.2. Time Allocation

Most of the work done by the participating women comprised activities traditionally considered to be "reproductive" work rather than "productive" work (Table 2) [117]. For most activity groups, the time allocations were not normally distributed (Supplementary Figure S1a–u). Based on the observation data, women spent over two-thirds of their time (median = 491 min) from 8 a.m. to 8 p.m. providing care. Domestic chores and own production were also important activities for allotted time (median = 318 min and 45 min, respectively). Overall, the highest amount of caregiving time was allotted to care of the index child (median = 405 min) and care of other children or adults (median = 255 min); for domestic chores, the highest amount of time overall was allotted to cooking (median = 85 min) and food preparation (median = 51 min) (Supplementary Table S12). At the population level, there was little time allocated to employment (median = 0 min). Among the women in this study who engaged in employed work (n = 77), 16 min (median) were allocated to that activity (Supplementary Table S4). Overall, although women spent about half their time socializing, hardly any time was allocated to other leisure activities (median = 0 min). Women in this study spend much of their time multitasking. The median number of concurrent activities across all timeslots was three (Supplementary Table S5). Women performed more than one activity in 88% of the 48 timeslots (Supplementary Table S6).

Table 2. Inter-method comparison of the median time allocated in minutes to activity groups. (Median value, and 25th and 75th percentiles).

ICATUS Activity Group	N	Non-Participation	OBS		24HR		IAR	
		n (%)	Median (minutes)	25th, 75th	Median (minutes)	25th, 75th	Median (minutes)	25th, 75th
Employment and related activities (MD1)	175	98 (56.0)	0	0, 5	0	0, 0	0	0, 35
Production of goods for own final use (MD2)	175	16 (9.1)	45	10, 79	49 †	15, 90	43	18, 81
Unpaid domestic services for household and family members (MD3)	175	0 (0.0)	318	263, 370	320	245, 396	311	251, 374
Unpaid caregiving services for household and family member (MD4) *	175	0 (0.0)	491	388, 608	180 †	96, 390	418 †	324, 541
Socializing and communication, community participation and religious practice (MD7) *	175	0 (0.0)	405	270, 525	195 †	75, 330	285 †	105, 465
Culture, leisure, mass media and sports practices (MD8) *	175	102 (58.3)	0	0, 30	0 †	0, 0	0 †	0, 0
Self-care and maintenance (MD9)	175	0 (0.0)	68	50, 88	58 †	39, 80	79 †	53, 111

OBS, observation; 24HR, 24 h recall; IAR, image-assisted recall. * Activity group contains one or more simultaneous activities. † *p*-value of Wilcoxon signed rank sum test of median time allocated compared to OBS < 0.05. NB: There were no activities that mapped to ICATUS MD5, unpaid volunteer, training, or other unpaid work. NB: This table does not include resting or sleeping (due to known inconsistencies with recording) or answering the phone for the study or other interactions for the study.

The inter-method comparisons show that the median amount of time allocated to caregiving was substantially underestimated by both 24HR (63%) and IAR (15%) (Table 2). Median time allocated to socializing was also substantially underestimated by both the 24HR (52%) and the IAR (30%) methods. Both 24HR and IAR methods accurately estimated the median time allocated to employment and domestic chores. The IAR method accurately estimated the median time allocated to own production whereas that median time was overestimated (9%) by the 24HR. The median time allocated to self-care was underestimated by the 24HR (15%) but overestimated by the IAR (16%). For most activities, median time allocations estimated via the EII (i.e., the enumerator's interpretation of the wearable camera's images compiled prior to the IAR) underestimated the observation data, ranging from 7% (domestic chores) to 78% (socializing) (Supplementary Table S3). The median number of concurrent activities was accurately estimated by the IAR but underestimated by the 24HR (Supplementary Table S5).

Comparing the median number of minutes estimated using the EII or IAR for the observation versus non-observation days showed no significant differences for employment, domestic chores, socializing (IAR only), leisure, and self-care (Supplementary Table S7a). However, for both EII and IAR, the median number of minutes allocated to own production on the observation day was lower than the non-observation day (35 min vs. 53 min for EII; 43 min vs. 60 min for IAR), whereas the median number of minutes allocated to caregiving was higher on the observation day than on the non-observation day (315 min vs. 235 min for EII; 418 min vs. 339 min for IAR). For the EII, the median time allocated socializing was also lower on the observation (90 min) than non-observation day (150 min).

3.3. Measures of Agreement

The systematic bias differs substantially across activity groups (Table 3). It is low for most activity groups (employment, own production, domestic chores, leisure, and self-care), ranging from 1 min (own production via IAR and domestic chores via 24HR) to 33 min (leisure via 24HR). However, for both methods the bias is high for caregiving (226 min for 24HR and 62 min for IAR) and socializing (172 min for 24HR and 109 min for IAR). For both 24HR and IAR, the percentage of participants with median time allocation estimations that were within 30 min of the criterion method ranged from 5% (caregiving via 24HR) to 79% (employment via 24HR) (Supplementary Table S8). Between 2% (self-care via 24HR) and 79% (caregiving via 24HR) of the time allocation estimates erred by more than two hours. For concurrent activities, there was no systematic bias for IAR whereas 24HR systematically underestimated the median number of concurrent activities by 1.3 (Supplementary Table S10). The difference in the estimated median number of concurrent activities (compared to observation) was less than two for about half (54%) of households via 24HR and three-quarters (74%) via IAR. (Supplementary Table S9). For only 17% and 21% of households (24HR and IAR, respectively), there was no difference in the estimated median number of concurrent activities compared to observation.

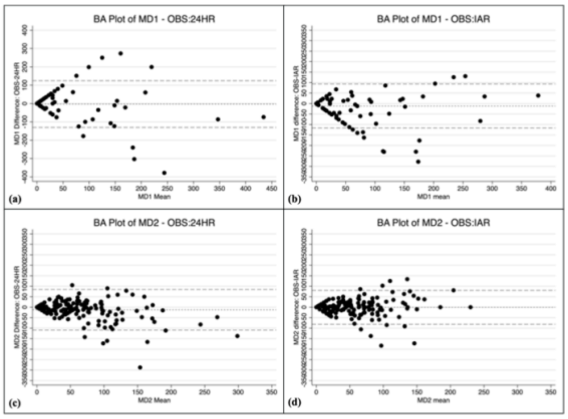
Table 3. Inter-method comparison of the time allocation bias and limits of agreement (LOA).

	Bias † (minutes)	**LOA ‡ minutes (hours)**	
Employment and related activities (MD1)			
24HR	−3	−130 (−2)	124 (2)
IAR	−12	−117 (−2)	94 (2)
Production of goods for own final use (MD2)			
24HR	−12	−109 (−2)	84 (1)
IAR	−1	−81 (−1)	80 (1)
Unpaid domestic services for household and family members (MD3)			
24HR	−1	−217 (−4)	215 (3)
IAR	8	−151 (−2)	167 (3)
Unpaid caregiving services for household and family members (MD4) *			

24HR	226	−223 (−4)	675 (11)
IAR	62	−267 (−4)	390 (7)
Socializing and communication, community participation, and religious practice (MD7) *			
24HR	172	−312 (−5)	656 (11)
IAR	109	−329 (−5)	548 (9)
Culture, leisure, mass media, and sports practices (MD8) *			
24HR	33	-169 (−3)	236 (4)
IAR	26	-189 (−3)	241 (4)
Self-care and maintenance (MD9)			
24HR	9	−73 (−1)	90 (2)
IAR	-17	−124 (−2)	90 (2)

LOA, limits of agreement; 24HR, 24 h recall; IAR, image-assisted recall. * Activity group contains one or more simultaneous activities. † Mean difference. ‡ +/− 2 SD from the mean difference. NB: A negative indicates that 24HR/IAR overestimated OBS.

For 24HR and IAR, the time allocation Bland–Altman plots showed varying patterns across activity groups (Figures 2a–n and 3a,b). Only the IAR method generated cloud-shaped plots (domestic chores, caregiving, and socializing), indicating the method performed equally well for women spending little time doing these activities and women spending substantial time doing these activities. Both the 24HR and the IAR methods had fan-shaped plots for employment, own production (IAR only), and leisure, indicating the amount of random error increased as the mean time allocated to the activity group increased. The 24HR method had several downward-sloping Bland–Altman plots (own production, domestic chores, caregiving, and self-care), and IAR had one downward-sloping plot for self-care, indicating the method underestimated time allocated to the activity for women on the lower end of the spectrum and overestimated time allocated to the activity for women at the upper end of the spectrum. For concurrent activities, the Bland–Altman plot appears cloud shaped for 24HR whereas the plot for IAR appears to be downward sloping.



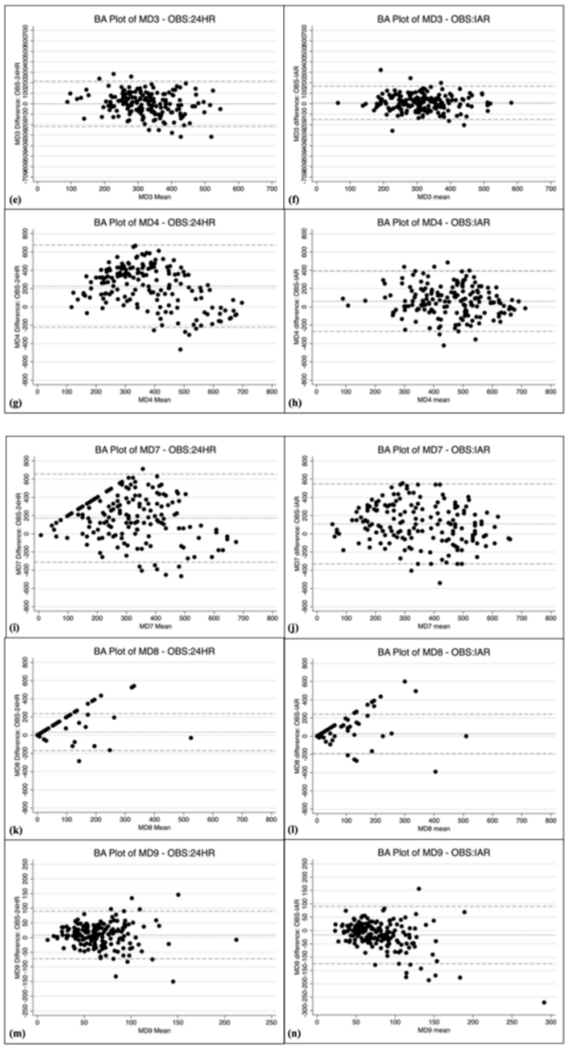


Figure 2. Bland–Altman (BA) plots of time allocation difference versus the mean of observation (OBS) and 24 h recall (24HR) or image assisted recall (IAR). (**a**) BA plot for employment and related activities (MD1), OBS and 24HR, (**b**) BA plot for employment and related activities (MD1), OBS and IAR, (**c**) BA plot for production of goods for own final use (MD2), OBS and 24HR and (**d**) BA plot for production of goods for own final use (MD2), OBS and IAR. (**e**) BA plot for unpaid domestic services for household and family members (MD3), OBS and 24HR, (**f**) BA plot for unpaid domestic

services for household and family members (MD3), OBS and IAR, (**g**) BA plot for unpaid caregiving services for household and family members (MD4), OBS and 24HR, and (**h**) BA plot for unpaid caregiving services for household and family members (MD4), OBS and IAR, (**i**) BA plot for socializing and communication, community participation, and religious practice (MD7), OBS and 24HR, (**j**) BA plot for socializing and communication, community participation, and religious practice (MD7), OBS and IAR, (**k**) BA plot for culture, leisure, mass media, and sports practices (MD8), OBS and 24HR, and (**l**) BA plot for culture, leisure, mass media, and sports practices (MD8), OBS and IAR, (**m**) BA plot for self-care and maintenance (MD9), OBS and 24HR, and (**n**) BA plot for self-care and maintenance (MD9), OBS and AR. The dotted line is the mean difference (bias), the long-dashed lines are +/− 2SD limits of agreement (LOA). A bias > 0 indicates that 24HR or IAR underestimates time allocation.

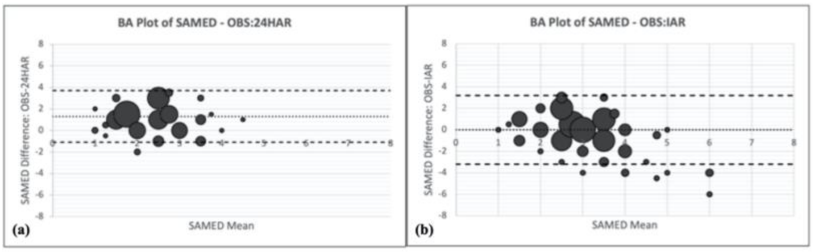


Figure 3. Bland–Altman (BA) plots of time allocation difference versus the mean of observation (OBS) and 24 h recall (24HR) or image-assisted recall (IAR). (**a**) BA plot for concurrent activities-OBS and 24HR, (**b**) BA plot for concurrent activities, OBS and IAR. The dotted line is the mean difference (bias), the long-dashed lines are +/− 2 SD limits of agreement (LOA). A bias > 0 indicates that 24HR or IAR underestimates time allocation. The size of the point corresponds with the number of households.

The width of the LOA varied substantially across activity groups (Table 3). The LOAs were within about 2 h for both methods for employment, own production, self-care, and domestic chores (IAR only). However, the LOAs for caregiving and socializing were high, with overestimates ranging from 223 to 329 min and underestimates ranging from 390 to 675 min. For concurrent activities, the LOA for the 24HR ranged from an overestimate of 1.1 activities to an underestimate of 3.7 activities, and for IAR +/− 3.2 activities (Supplementary Table S10).

For both the 24HR and IAR methods, Cronbach's coefficient alpha indicated that the inter-method agreement with observation was unacceptable for most activities (caregiving, socializing, leisure, and self-care) (Table 4). For domestic chores, the reliability was also unacceptable for 24HR but was acceptable for IAR. For own production, the reliability was moderate for 24HR but acceptable for IAR. Reliability for employment for both methods was acceptable. For concurrent activities, Cohen's κ indicated that agreement was no better than if it had occurred purely by chance (24HR = 0.028; IAR = 0.031) (Supplementary Table S11).

Table 4. Inter-method comparison of reliability for time allocation.

ICATUS Activity Group	24HR		IAR	
	alpha	Score †	alpha	Score †
Employment and related activities (MD1)	0.7347	acceptable	0.7847	acceptable
Production of goods for own final use (MD2)	0.8056	moderate	0.7938	acceptable
Unpaid domestic services for household and family members (MD3)	0.6014	unacceptable	0.7618	acceptable
Unpaid caregiving services for household and family members (MD4) *	0.2901	unacceptable	0.4273	unacceptable

Socializing and communication, community participation, and religious practice (MD7) *	0.1728	unacceptable	0.2270	unacceptable
Culture, leisure, mass media, and sports practices (MD8) *	0.5107	unacceptable	0.3881	unacceptable
Self-care and maintenance (MD9)	0.4455	unacceptable	0.3792	unacceptable

24HR, 24 h recall; IAR, image-assisted recall; CI, confidence interval. * Activity group contains one or more simultaneous activities. † Using Cronbach's (reliability) coefficient alpha. Nunnally (1978) and Peterson (1994) suggest the following benchmark scale for interpreting the alpha statistic: <0.70 unacceptable; >0.70 acceptable; >0.80 moderate; 0.90–0.95 high; >0.95 suspect.

4. Discussion

This is the first study to validate the 24HR or IAR methods using an AWC for collecting women's time-use data in a low-income country context. We assessed the concurrent validity using direct observation as the criterion method with 211 women in the rural Eastern Region of Uganda. The results show the systematic bias for time allocation to employment, own production, domestic chores, and self-care was low, for both the 24HR and IAR, whereas time allocation to caregiving and socializing may be severely underestimated (>1 h) by both methods. The extent of underestimation at the population level was higher for the 24HR than IAR, especially for caregiving (3.5 times higher). This finding is consistent with other studies, which show between a third and three-quarters of respondents' recall regarding childcare is inaccurate [69], and that, compared with other activities recall errors are highest for caregiving activities such as feeding children, breastfeeding, and supervising children [50].

Several factors may have contributed to the systematic underestimation of time allocation to caregiving and socializing seen in this study. First, most childcare in this study context (rural Uganda) is omnipresent "passive" childcare, that is, constantly performed while simultaneously performing other household chores or chatting, e.g., a mother may supervise a small child at play while washing clothes. Such omnipresent "passive" activities may be so routine as to seem unremarkable to the participant in both the AWC photographic record and in memory [118]. Second, some activities, such as socializing, largely happen "off-camera", e.g., while washing clothes, the mother may be chatting with a friend who is not in the camera's field of vision. There is no photographic record of these "background" activities to trigger the participant's recall. In both examples, the AWC photos used in the IAR may remediate some but not all misreporting which could explain why the LOAs are wider and the systematic bias is higher for the 24HR than IAR. Further support for this interpretation is the numbers of concurrent activities that women performed were higher and accurately estimated by IAR whereas they were underestimated by the 24HR.

This finding is also consistent with results from previous studies showing, if concurrent activities were taken into account, estimates of women's time allocation to childcare would increase two-fold [119], and that less than a quarter of time spent on childcare is reported via traditional methods as a "primary" activity [34]. In a multiple-country analysis, estimates of women's time allocation to childcare increased depending on how concurrent activities were counted, ranging from an increase of 31% (Ethiopia) to 134% (Zimbabwe) [120]. There is a long history of discussion on how to count time allocated to concurrent activities [20,34,43,121]. Most studies avoid dealing with multiple concurrent activities by artificially limiting the number of activities collected or analyzed (e.g., just the "primary" activity). When multiple concurrent activities are allowed, typically the timeslot is equally divided among the concurrent activities, which presumes these activities are performed sequentially [122]. This presumption, however, does not hold true in a rural Uganda setting where childcare or chatting are generally done concurrently with other activities. In this study, if caregiving and socializing had been analyzed in the traditional way (i.e., treating them as sequential rather than simultaneous activities), the estimated median amount of time allocated to these activities would have been reduced by

44% and 66%, respectively (Supplementary Table S13). In research where women's time use is an outcome of interest, the method of data collection and analysis must account for both concurrent and simultaneous activities to accurately reflect women's time burdens and social well-being.

For most activities, random error could be high (greater than 2 h), most notably for caregiving and socializing where underestimates could exceed 7 h. Such high LOAs indicate that at the individual level, for most activities, inaccuracies in time estimations can be large using either the 24HR or IAR. If results from 24HR or IAR are used to assess their associations with other variables, attenuation will occur. The finding of large random errors is consistent with other time-use studies, although the cause is unclear [89,123].

For time allocated to domestic chores, caregiving, and socializing, only the IAR showed a Bland–Altman plot pattern having no slope, indicating constant variability of the error. Therefore, compared to 24HR, using the IAR for these activities may result in more predictable bias when using time use as predictor of an outcome variable in regression outcome models. There was a downward slope in the Bland–Altman plots for domestic chores and caregiving (24HR only). Fan-shaped plot patterns were also found (employment, leisure, socializing for 24HR only, and own production for IAR only). Fan-shaped and downward-sloping plot patterns indicate that, for these activities and assessment methods, the magnitude and/or direction of measurement error may change as the amount of time performing the activity increases. This precludes any attempt to predict the consequences of the measurement error on regression outcome models including time use as an exposure/predictor.

Several AWC studies for other outcomes of interest (diet, physical activity, caregiving) did not include an IAR. Instead, a topical expert coded the images based on their interpretation of the activities recorded in the photos, which reduces respondent burden [71,73,75,87,89,95,96,116,124–126]. We therefore examined whether the IAR was essential for the interpretation of the AWC photos. In this study, the EII did not provide a reliable estimate of women's time allocation. The pattern was the same as the IAR but the degree of underestimation compared to direct observation was more severe. For example, the EII underestimated the median number of minutes allocated to caregiving and socializing by 37% and 78%, respectively, compared with 17% and 30%, for the IAR. These results indicate insufficient visual clues were captured for an external coder to determine all activities undertaken in a rural low-income country setting where women primarily work from home.

Several studies have investigated the associations between women's time allocations and maternal and/or child nutrition-related outcomes [8,11–18,127,128]. The results of these studies, which are based on 24HRs, are often conflicting. The results of this study indicate that previous time-use research relying on the 24HR method, in particular research exploring the associations between women's time allocations and nutrition outcomes, is likely unreliable.

Strengths and Limitations

One strength of this study is the use of direct observation as the criterion method. The process of observation, however, might have influenced the participants' activities or IAR proficiency (e.g., due to heightened awareness of activities performed). Comparisons of time allocation on observation day vs. non-observation day measured via IAR indicate that study participants recalled more time caregiving and less time engaged in own production on observation days than non-observation days. The same is true of the EII, suggesting that the difference in time allocation is real and not just a difference in recall (Supplementary Table S7a). More time was spent caregiving on observation days compared to non-observation days regardless of the order in which the household was observed versus administered the IAR method (Supplementary Table S7b,c). Even though the increased time spent caregiving might be due to a social desirability or reactivity bias, it may simply reflect a culture of hospitality. The median time spent caregiving remained high on non-observation days (339 min vs. 418 min observed day) and well above that assessed via 24HR on the observation day (180 min), indicating that any changes in activity patterns due to having an observer at home did not substantially contribute to under-reporting of caregiving activities.

The quality of data captured by the AWC was often compromised by technical issues that have also been reported by previous investigators. These issues include insufficient lighting or poor image quality [75,80,84–87,89–92,98,126,129–131]; the tedious, time-consuming, and manual processes required to manage and code hundreds of thousands of photos [75,80,88–90,98,124]; hardware issues resulting in lost data [71,73,75–77,80,81,83,89–91,95,98,124,126,131–134]; and camera-specific software issues (e.g., the built-in filename format by image number rather than timestamp, tendency of the cameras to "lose" time over time). Aside from the functionality of the AWCs, the onerous structure of the IAR protocol for enumerators may have contributed to error. The observation and 24HR protocols were closed-ended and shared a similar matrix structure of pre-specified activity categories, whereas the IAR structure was open-ended to capture a narrative of the activities performed to elicit more detailed information. Some of these difficulties could be remediated in the future with a computer-assisted personal interview (CAPI)-based data collection tool that could prompt enumerators when, for example, a series of recorded activities was not closed.

Although all enumerators and research assistants received the same training in the coding of activities, there were variations in how the same or similar activities were coded between field enumerators and also between IAR data entry research assistants. Inconsistencies in the coding of activities within activity groups posed no problem, since the time allocation was analyzed at the activity group level. For example, playing with a young child and feeding a young child are somewhat ambiguous activities, but both fell under the caregiving activity group. Underestimations in one activity would have offset overestimates in another activity within the same activity group. However, a few activities were coded into different activity groups by field enumerators and data entry research assistants. For example, peeling sweet potatoes and shelling and pounding groundnuts were variably coded as food preparation (domestic chores) or post-harvest processing (own production). To address this issue, post-harvest processing activities were re-mapped to the domestic chores activity group. While this issue was caught, other coding inconsistencies may have contributed to the high random error seen in this study.

Due, in part, to the poor performance of the AWC in the low-light conditions of early morning and late evening, we restricted the analysis (for all three methods) from the intended 15 h period to a standard 12 h period (8 a.m. to 8 p.m.). Restricting the period of analysis may have influenced the results if one method was better than the other at capturing an activity that occurred primarily outside the 8 a.m. to 8 p.m. period. Upon review (ALSB), however, the only activities that commonly occurred before 8 a.m. or after 8 p.m.—and not any other part of the day—were study-related interactions, which were not included in the analysis. Whereas cooking food and eating often happened after 8 p.m., these activities also always occurred during the day too, so any differences in method performance would be evident. Finally, less than half of the women in this study were engaged in employed work and the data were only collected in one season. The results of the inter-method comparisons may be different in a different season or in a population where a larger proportion of women spent more time in employed work if, for example, one and/or the other method was more effective at measuring time allocated to employment.

5. Conclusions

This study aimed to evaluate the concurrent criterion validity, for assessing women's time use, using an AWC-based IAR method and the 24HR method. Our hypothesis was that prospectively capturing activity data would reduce systematic and random errors inherent to time allocation recalls and reduce respondent/interviewer burdens inherent to observation to allow accurate time allocation data collection at scale for programmatic purposes in rural low-income country contexts. Our results indicate that both the 24HR and IAR provide accurate estimates of the median time allocated to employment, own production (IAR only), and domestic chores at the population level, whereas neither the 24HR nor IAR are valid methods for measuring median time allocated to caregiving or socializing. The high LOAs observed across all activities indicate high random error at the individual level, which will attenuate true associations between time allocation—where estimated via 24HR or IAR—and outcomes of interest. For most activities, neither 24HR nor IAR are valid methods for estimating time allocation at the individual level. To the best of our knowledge, there is no globally accepted threshold LOA for time use, however, a difference of more than two hours (out of twelve hours) seems substantial. The cloud-shaped pattern exhibited only by the IAR-generated Bland–Altman plots for own production, domestic chores, caregiving, and socializing suggest that measurement error due to IAR may be easier to handle and adjust for statistically compared to 24HR when assessing associations of time use with outcome variables for these activities.

This study has important implications for interpreting time-use data collected via 24HR in, for example, the Women's Empowerment in Agriculture Index (WEAI) standard time-use module [65] or the Living Standards Measurement Study (LSMS) stylized activity log module [135]. It suggests concurrent activities, such as socializing and caregiving, may be under-estimated unless explicitly probed and counted. These results lend credence to modifications made in the time allocation module of the Project-Level Women's Empowerment in Agriculture Index (Pro-WEAI) to reduce error in measuring caregiving with the addition of checkboxes for each 24HR timeslot indicating whether the participant was caring for a child [136,137]. The same approach may be needed for time spent socializing.

In calculating time allocation estimates for these activities, they should be allotted credit for the entire timeslot duration. Formative research conducted in the study location is also important to understand the activities commonly undertaken by the target population, their purpose, in terms of own use or income generation, and patterns (simultaneous or not) so that they can be properly categorized.

This study shows that, for caregiving, socializing, and domestic chores, the IAR outperforms 24HR. This is important because caregiving and domestic chores are activities most often performed by women. Further work is needed to design an IAR protocol that works in rural low-income country contexts where literacy is low and exposure to first-person perspective photographs is limited. The IAR protocol should be simplified and modified to enable image coding in the field. Instead of reviewing all images with the respondent, it may be more practical and effective to probe the respondent on activities using a few pre-selected (by the enumerator) "sentinel" images per timeslot. Furthermore, enumerator training should include practice recognizing and interpreting problematic activities using AWC photos collected from target populations in the study area, such as breastfeeding and passive caregiving; scanning for contextual clues in individual images and across a series of images; and facilitation skills. Coding consistency across enumerators should be assessed prior to the start of data collection.

Further research is needed to understand how low-literacy populations with limited exposure to first-person perspective photographs cognitively process wearable camera images. The IAR method assumes that AWC-generated photographs will trigger the participants' memory of activities done on the previous day to improve recall accuracy [100–103]. If the participants inferred what they were doing from what they saw in the photos rather than used the photos as a memory aid, it would be a very different cognitive task and (possibly) outcome. As far as we are aware, this is the first study to quantify the extent of measurement error, when the 24HR or AWC-IAR are used to estimate women's time use in a low-income country context. Future research should also assess the magnitude and nature of error in estimating time allocation with 24HR and IAR in other contexts.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/nu14091833/s1, Figure S1: Plots of the quantiles of activities against the quantiles of the normal distribution (Q–Q plot); Table S1: Household data collection patterns; Table S2: Time-use activities and ICATUS major divisions; Table S2: Time-use activities and ICATUS major divisions; Table S3: Inter-method comparison of the median time allocated in minutes to activity groups, including enumerator image interpretation (EII); Table S4: Inter-method comparison of the median time allocated in minutes to activity groups for participating women only; Table S5: Inter-method comparison of the median number of concurrent activities; Table S6: Inter-method comparison of the median number and proportion of timeslots containing concurrent activities; Table S7a: Inter-method comparison of the median time allocated in minutes to activity groups, observation day versus non-observation day; Table S7b: Inter-method comparison of the median time allocated in minutes to activity groups, observation day versus non-observation day for households having IAR administered before OBS only; Table S7c: Inter-method comparison of the median time allocated in minutes to activity groups, observation day versus non-observation day for households having IAR administered after OBS only; Table S8: Frequency of inter-method time allocation differences; Table S9: Frequency of inter-method median concurrent activities differences; Table S10: Inter-method comparison of the median concurrent activities bias and limits of agreement (LOA); Table S11: Inter-method comparison of reliability for median concurrent activities; Table S12: Inter-method comparison of the median time allocated in minutes to discrete activities; Table S13: Within-method comparison of time allocation—with and without simultaneous activities.

Author Contributions: Conceptualization, L.P., J.P., J.N., K.W. and E.L.F.; methodology, A.L.S.B., L.P., J.P., L.F., G.V., J.N., K.W. and E.L.F.; validation, A.L.S.B., L.P. and E.L.F.; formal analysis, A.L.S.B., L.P. and E.L.F.; investigation, A.L.S.B., G.V. and J.N.; data curation, A.L.S.B., G.V., B.D.G. and N.S.; writing—original draft preparation, A.L.S.B.; writing—review and editing, L.P., L.F., P.K., G.V., B.D.G., N.S., J.N., K.W. and E.L.F.; visualization, A.L.S.B.; supervision, L.P., J.N., K.W. and E.L.F.; project administration, J.N., K.W. and E.L.F.; funding acquisition, L.P., J.P., J.N., K.W. and E.L.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by an Innovative Methods and Metrics for Agriculture and Nutrition Actions (IMMANA) grant, funded by UK AID, grant number IMMANA 2.05.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Uganda National Council for Science and Technology (UNCST) (A24ES), the London School of Hygiene & Tropical Medicine Observational/Interventions Research Ethics Committee (Project ID: 1420), and the University of Greenwich Faculty of Engineering and Science Ethics Committee (Project ID: B0501).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data that support the findings of this study are openly available in Dataverse at (URL), reference number (reference number).

Acknowledgments: We acknowledge the contributions of the field team and thank the mothers who participated in this study, and their families and communities.

Conflicts of Interest: The authors declare no conflict of interest. IMMANA had no role in the design, analysis, or writing of this article.

References

1. Engle, P.L.; Menon, P.; Haddad, L. *Care and Nutrition: Concepts and Measurement*; International Food Policy Research Institute: Washington, DC, USA, 1997.
2. Blackden, C.M.; Wodon, Q. *Gender, Time Use, and Poverty in Sub-Saharan Africa*; World Bank Working Papers; The World Bank: Washington, DC, USA, 2006.
3. OECD. *Enabling Women's Economic Empowerment: New Approaches to Unpaid Care Work in Developing Countries*; OECD Publishing: Paris, France, 2019.
4. International Labour Office. *Women at Work: Trends 2016*; International Labour Office: Geneva, Switzerland, 2016.
5. Verick, S. Female labor force participation in developing countries. *IZA World Labor* **2014**, *2014*, 87**.**
6. Marini, M.; Shelton, B. Measuring Household Work: Recent Experience in the United States. *Soc. Sci. Res.* **1993**, *22*, 361–382.
7. Kadiyala, S.; Harris, J.; Headey, D.; Yosef, S.; Gillespie, S. Agriculture and nutrition in India: Mapping evidence to pathways. *Ann. N. Y. Acad. Sci.* **2014**, *1331*, 43–56.

8. Komatsu, H.; Malapit, H.J.L.; Theis, S. Does women's time in domestic work and agriculture affect women's and children's dietary diversity? Evidence from Bangladesh, Nepal, Cambodia, Ghana, and Mozambique. *Food Policy* **2018**, *79*, 256–270.
9. Molyneux, M. Mothers at the Service of the New Poverty Agenda: Progresa/Oportunidades, Mexico's Conditional Transfer Programme. *Soc. Policy Adm.* **2006**, *40*, 425–449.
10. Senauer, B. *The Impact of the Value of Women's Time on Food and Nutrition in Developing Countries*; University of Minnesota: Minneapolis, MN, USA, 1988.
11. Jones, A.D.; Agudo, Y.C.; Galway, L.; Bentley, J.; Pinstrup-Andersen, P. Heavy agricultural workloads and low crop diversity are strong barriers to improving child feeding practices in the Bolivian Andes. *Soc. Sci. Med.* **2012**, *75*, 1673–1684.
12. van den Bold, M.; Bliznashka, L.; Ramani, G.; Olney, D.; Quisumbing, A.; Pedehombga, A.; Ouedraogo, M. Nutrition-sensitive agriculture programme impacts on time use and associations with nutrition outcomes. *Matern. Child Nutr*. **2021**, *17*, 16.
13. Ricci, J.A.; Jerome, N.W.; Sirageldin, I.; Aly, H.; Moussa, W.; Galal, O.; Harrison, G.G.; Kirksey, A. The significance of children's age in estimating the effect of maternal time use on children's well-being. *Soc. Sci. Med.* **1996**, *42*, 651–659.
14. Popkin, B.M. Time allocation of the mother and child nutrition. *Ecol. Food Nutr.* **1980**, *9*, 1–13.
15. Gryboski, K.L. Maternal and Non-maternal Time Allocation to Infant Care, and Care during Infant Illness in Rural Java, Indonesia. *Soc. Sci. Med.* **1996**, *43*, 209–219.
16. Hawkes, K.; O'Connell, J.F.; Jones, N.G.B. Hadza Women's Time Allocation, Offspring Provisioning, and the Evolution of Long Postmenopausal Life Spans. *Curr. Anthr.* **1997**, *38*, 551–577.
17. Nordang, S.; Shoo, T.; Holmboe-Ottesen, G.; Kinabo, J.; Wandel, M. Women's work in farming, child feeding practices and nutritional status among under-five children in rural Rukwa, Tanzania. *Br. J. Nutr.* **2015**, *114*, 1594–1603.
18. Bamji, M.S.; Thimayamma, B. Impact of women's work on maternal and child nutrition. *Ecol. Food Nutr.* **2000**, *39*, 13–31.
19. Hirway, I.; Jose, S. Understanding Women's Work Using Time-Use Statistics: The Case of India. *Fem. Econ.* **2011**, *17*, 67–92.
20. Fedick, C.B.; Pacholok, S.; Gauthier, A.H. Methodological issues in the estimation of parental time–Analysis of measures in a Canadian time-use survey. *Electron. Int. J. Time Use Res.* **2005**, *2*, 67–87.
21. Buvinić, M.; King, E.M. Invisible No More? A Methodology and Policy Review of How Time Use Surveys Measure Unpaid Work. United Nations Foundation, Data2X. Available online: https://data2x.org/resource-center/invisible-no-more-a-methodology-and-policy-review-of-how-time-use-surveys-measure-unpaid-work-volume-1/ (accessed on 30 March 2022).
22. Bério, A.-J. The analysis of time allocation and activity patterns in nutrition and rural development planning. *Food Nutr. Bull.* **1984**, *6*, 1–16.
23. Walthery, P.; Gershuny, J. Improving Stylised Working Time Estimates with Time Diary Data: A Multi Study Assessment for the UK. *Soc. Indic. Res.* **2019**, *144*, 1303–1321.
24. Kanwar, P.; Yadav, D.; Sharma, N. Time use pattern of hill farm women: A study in Himachal Pradesh. *Himachal. J. Agric. Res.* **2003**, *29*, 84–88.
25. Szalai, A. Women's time: Women in the light of contemporary time-budget research. *Futures* **1975**, *7*, 385–399.
26. Engle, P.; Lhotská, L.; Armstrong, H.; UNICEF. *The Care Initiative: Assessment, Analysis and Action to Improve Care for Nutrition*; UNICEF: New York, NY, USA, 1997; p. 35.
27. Floro, M.S.; King, E.M. The present and future of time-use analysis in developing countries. *Asia-Pac. Popul. J.* **2016**, *31*, 5–42.
28. Johnston, D.; Stevano, S.; Malapit, H.J.L.; Hull, E.; Kadiyala, S. *Agriculture, Gendered Time Use, and Nutritional Outcomes: A Systematic Review*; Report No.: 01456; IFPRI Discussion Paper; The International Food Policy Research Institute (IFPRI): Washington, DC, USA, 2015.
29. Mcguire, J.S.; Popkin, B.M. *Helping Women Improve Nutrition in the Developing World*; The International Bank for Reconstruction and Development: Washington, DC, USA; The World Bank: Washington, DC, USA, 1990.
30. Floro, M.S. Economic restructuring, gender and the allocation of time. *World Dev.* **1995**, *23*, 1913–1929.
31. Nti, C.A.; Inkumsah, D.; Fleischer, G. Influence of women's workload on their nutritional status in selected communities in Ghana. *J. Consum. Stud. Home Econ.* **1999**, *23*, 165–170.
32. Leslie, J. Women's time: A factor in the use of child survival technologies? *Health Policy Plan* **1989**, *4*, 1–16.
33. Leslie, J.; Paolisso, M.J. *Women, Work, and Child Welfare in the Third World*; AAAS selected symposia series; Westview Press for the American Association for the Advancement of Science: Washington, DC, USA, 1989; p. 265.
34. Ironmonger, D. There are only 24 Hours in a Day! Solving the problematic of simultaneous time. In Proceedings of the 25th IATUR Conference on Time Use Research, Lisbon, Portugal, 16–18 October 2005; p. 8.
35. Irani, L.; Vemireddy, V. Getting the measurement right! Quantifying time poverty and multitasking from childcare among mothers with children across different age groups in rural north India. *Asian Popul. Stud.* **2020**, *17*, 94–116.
36. Peterman, A.; Ng, S.W.; Palermo, T.; Lee, I.-H.E. Managing the Double Burden: Pregnancy and Labor-Intensive Time Use in Rural China, Mexico, and Tanzania. *Stud. Fam. Plan.* **2013**, *44*, 411–430.
37. Mattingly, M.J.; Blanchi, S.M. Gender Differences in the Quantity and Quality of Free Time: The, U.S. Experience. *Soc. Forces* **2003**, *81*, 999–1030.
38. Zick, C.D.; Bryant, W. A New Look at Parents' Time Spent in Child Care: Primary and Secondary Time Use. *Soc. Sci. Res.* **1996**, *25*, 260–280.
39. Floro, M.S. Women's well-being, poverty, and work intensity. *Fem. Econ.* **1995**, *1*, 1–25.
40. Ilahi, N. *The Intra-household Allocation of Time and Tasks: What Have We Learnt from the Empirical Literature? Policy Research Report on Gender and Development Working Paper*; The World Bank: Washington, DC, USA, 2000.

41. Acharya, M.; Bennett, L. *Women and the Subsistence Sector: Economic Participation and Household Decisionmaking in Nepal*; Report No.: 526; World Bank Staff Working Papers: Washington, DC, USA, 1982.
42. As, D. Studies of Time-Use: Problems and Prospects. *Acta. Sociol.* **1978**, *21*, 125–141.
43. Budlender, D. *A Critical Review of Selected Time Use Surveys*; Report No.: 2; Gender and Development Programme Paper; United Nations Research Institute for Social Development: Geneva, Switzerland, 2007.
44. Budlender, D. *The Statistical Evidence on Care and Non-Care Work across Six Countries*; Report No.: 4; Gender and Development Programme Paper; United NationsResearch Institutefor Social Development: Geneva, Switzerland, 2008.
45. ILO Department of Statistics. *Survey Methods to Improve Measurement of Paid and Unpaid Work: Country Practices in Time-Use Measurement*; ILO Department of Statistics: Geneva, Switzerland, 2018; p. 40.
46. Andorka, R. Time budgets and their uses. *Annu. Rev. Sociol.* **1987**, *13*, 149–164.
47. European Union. *Harmonised European Time Use Surveys (HETUS) 2018 Guidelines*; Publications Office of the European Union: Luxembourg, 2019.
48. Michelson, W. Time Budgets and Social Activity. In *Meeting of the Working Group on Time-Budgets and Social Activity*; Major Report No. 4; University of Toronto: Toronto, ON, Canada 1975; Volume 1, p. 213.
49. Hirway, I. Time-Use Surveys in Developing Countries: An Assessment. In *Unpaid Work and the Economy*; Palgrave Macmillan: London, UK, 2010; pp. 252–324.
50. Ricci, J.; Jerome, N.; Megally, N.; Galal, O.; Harrison, G.; Kirksey, A. Assessing the Validity of Informant Recall: Results of a Time Use Pilot Study in Peri-Urban Egypt. *Hum. Organ.* **1995**, *54*, 304–308.
51. Robinson, J.P. The time-diary method: Structure and uses. In *Time Use in the Social Sciences*; Pentland, W.E., Harvey, A.S., Lawton, M.P., McColl, M.A., Eds.; Kluwer Academic Publishers: New York, NY, USA, 1999; pp. 47–89.
52. Gross, D.R. Time allocation: A tool for the study of cultural behavior. *Annu. Rev. Anthropol.* **1984**, *13*, 519–558.
53. Kes, A.; Swaminathan, H. Gender and Time Poverty in Sub-Saharan Africa. In *Time Use, and Poverty in Sub-Saharan Africa*; Blackden, C.M., Wodon, Q., Eds.; The International Bank for Reconstruction and Development: Washington, DC, USA; The World Bank: Washington, DC, USA, 2006; pp. 13–38.
54. Kalton, G. Sample Design Issues in Time Diary Studies. In *Time, Goods, and Well-Being*; Juster, F.T., Stafford, F.P., Eds.; University of Michigan: Ann Arbor, MI, USA, 1985; pp. 93–112.
55. Robinson, J.P. The Validity and Reliability of Diaries Versus Alternative Time Use Measures. In *Time, Goods, and Well-Being*; Juster, F.T., Stafford, F.P., Eds.; The University of Michigan: Ann Arbor, MI, USA, 1985; pp. 33–62.
56. Wagenbuur, H.T.M. *An Analysis of the Time-Budget and of the Production and Productivity of Lime Farmers in Southern Ghana*; I.S.S. Ocassional Papers; Institute of Social Studies: The Hague, The Netherlands, 1972.
57. Gardner, F. Methodological Issues in the Direct Observation of Parent–Child Interaction: Do Observational Findings Reflect the Natural Behavior of Participants? *Clin. Child Fam. Psychol. Rev.* **2000**, *3*, 185–198.
58. Repp, A.C.; Nieminen, G.S.; Olinger, E.; Brusca, R. Direct Observation: Factors Affecting the Accuracy of Observers. *Counc. Rev.* **1988**, *55*, 29–36.
59. Richer, J. Direct Observation: Impediments and Approaches. *Hum. Ethol.* **2017**, *32*, 6–14.
60. Seymour, G.; Malapit, H.; Quisumbing, A. *Measuring Time Use in Development Settings*; International Food Policy Research Institute (IFPRI): Washington, DC, USA, 2016.
61. Juster, F.T.; Ono, H.; Stafford, F.P. An Assessment of Alternative Measures of Time Use. *Sociol. Methodol.* **2003**, *33*, 19–54.
62. Brunnich, B.; Druce, P.; Ghissassi, M.; Johnson, M.; Majidi, N.; Radas, A.L.; Riccheri, P.R.; Camille de Sentenac, C.; Vacarr, D. *Three Case Studies of Time Use Survey Application in Lower and Middle-Income Countries*; The Gender Unit in the Bureau for Development Policy at The United Nations Development Programme: Paris, France, 2005.
63. Belli, R.F.; Schwarz, N.; Singer, E.; Talarico, J. Decomposition can harm the accuracy of behavioural frequency reports. *Appl. Cogn. Psychol.* **2000**, *14*, 295–308.
64. Seymour, G.; Malapit, H.; Quisumbing, A. Measuring Time Use in Developing Country Agriculture: Evidence from Bangladesh and Uganda. *Fem. Econ.* **2020**, *26*, 169–199.
65. Alkire, S.; Meinzen-Dick, R.; Peterman, A.; Quisumbing, A.R.; Seymour, G. *The Women's Empowerment in Agriculture Index*; OPHI Working Paper; Report No.: 58; Oxford Poverty & Human Development Initiative: Oxford, UK, 2013.
66. Rubin, D.C.; Baddeley, A.D. Telescoping is not time compression: A model. *Mem. Cogn.* **1989**, *17*, 653–661.
67. Tourangeau, R. Cognitive Sciences and Survey Methods. In *Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines*; Jabine, T., Straf, M., Tanur, J., Tourangeau, R., Eds.; National Academy Press: Washington, DC, USA, 1984; pp. 73–100.
68. Jacobs, J.A. Measuring time at work: Are self-reports accurate? *Mon. Lab. Rev.* **1998**, *121*, 42–53.
69. Bernard, H.R.; Killworth, P.; Kronenfeld, D.; Sailer, L. The Problem of Informant Accuracy: The Validity of Retrospective Data. *Annu. Rev. Anthropol.* **1984**, *13*, 495–517.
70. United Nations Statistics Division. *International Classification of Activities for Time-Use Statistics 2016 (ICATUS 2016)*; United Nations Statistics Division: New York, NY, USA, 2019.
71. Arab, L.; Winter, A. Automated Camera-Phone Experience with the Frequency of Imaging Necessary to Capture Diet. *J. Am. Diet. Assoc.* **2010**, *110*, 1238–1241.
72. Arab, L.; Estrin, D.; Kim, D.H.; Burke, J.; Goldman, J. Feasibility testing of an automated image-capture method to aid dietary recall. *Eur. J. Clin. Nutr.* **2011**, *65*, 1156–1162.

73. Beltran, A.; Dadabhoy, H.; Ryan, C.; Dholakia, R.; Jia, W.; Baranowski, J.; Sun, M.; Baranowski, T. Dietary Assessment with a Wearable Camera among Children: Feasibility and Intercoder Reliability. *J. Acad. Nutr. Diet.* **2018**, *118*, 2144–2153.
74. Kamar, M.; Evans, C.; Hugh-Jones, S. Factors Influencing British Adolescents' Intake of Whole Grains: A Pilot Feasibility Study Using SenseCam Assisted Interviews. *Nutrients* **2019**, *11*, 2620.
75. Chan, V.; Davies, A.; Wellard-Cole, L.; Lu, S.; Ng, H.; Tsoi, L.; Tiscia, A.; Signal, L.; Rangan, A.; Gemming, L.; et al. Using Wearable Cameras to Assess Foods and Beverages Omitted in 24 Hour Dietary Recalls and a Text Entry Food Record App. *Nutrients* **2021**, *13*, 1806.
76. Gemming, L.; Doherty, A.; Kelly, P.; Utter, J.; Ni Mhurchu, C. Feasibility of a SenseCam-assisted 24-h recall to reduce under-reporting of energy intake. *Eur. J. Clin. Nutr.* **2013**, *67*, 1095–1099.
77. O'Loughlin, G.; Cullen, S.J.; McGoldrick, A.; Connor, S.; Blain, R.; O'Malley, S.; Warrington, G. Using a Wearable Camera to Increase the Accuracy of Dietary Analysis. *Am. J. Prev. Med.* **2013**, *44*, 297–301.
78. Jia, W.; Chen, H.-C.; Yue, Y.; Li, Z.; Fernstrom, J.; Bai, Y.; Li, C.; Sun, M. Accuracy of food portion size estimation from digital pictures acquired by a chest-worn camera. *Public Health Nutr.* **2013**, *17*, 1671–1681.
79. Gemming, L. *Image-Assisted Dietary Assessment: Evaluating the Potential of Wearable Cameras to Enhance Self-Report in the 24-Hour Dietary Recall Method*; The University of Auckland: Auckland, New Zealand, 2015.
80. Gemming, L.; Doherty, A.; Utter, J.; Shields, E.; Ni Mhurchu, C. The use of a wearable camera to capture and categorise the environmental and social context of self-identified eating episodes. *Appetite* **2015**, *92*, 118–125.
81. Gemming, L.; Rush, E.; Maddison, R.; Doherty, A.R.; Gant, N.; Utter, J.; Ni Mhurchu, C. Wearable cameras can reduce dietary under-reporting: Doubly labelled water validation of a camera-assisted 24 h recall. *Br. J. Nutr.* **2014**, *113*, 284–291.
82. Gemming, L.; Ni Mhurchu, C. Dietary under-reporting: What foods and which meals are typically under-reported? *Eur. J. Clin. Nutr.* **2016**, *70*, 640–641.
83. Pettitt, C.; Liu, J.; Kwasnicki, R.M.; Yang, G.-Z.; Preston, T.; Frost, G. A pilot study to determine whether using a lightweight, wearable micro-camera improves dietary assessment accuracy and offers information on macronutrients and eating rate. *Br. J. Nutr.* **2015**, *115*, 160–167.
84. Kelly, P.; Doherty, A.; Berry, E.; Hodges, S.; Batterham, A.M.; Foster, C. Can we use digital life-log images to investigate active and sedentary travel behaviour? Results from a pilot study. *Int. J. Behav. Nutr. Phys. Act.* **2011**, *8*, 44.
85. Kelly, P.; Doherty, A.R.; Hamilton, A.; Matthews, A.; Batterham, A.M.; Nelson, M.; Foster, C.; Cowburn, G. Evaluating the Feasibility of Measuring Travel to School Using a Wearable Camera. *Am. J. Prev. Med.* **2012**, *43*, 546–550.
86. Kelly, P. Assessing the Utility of Wearable Cameras in the Measurement of Walking and Cycling. Ph.D. Thesis, Oxford University, Oxford, UK, 2013.
87. Kerr, J.; Marshall, S.J.; Godbole, S.; Chen, J.; Legge, A.; Doherty, A.R.; Kelly, P.; Oliver, M.; Badland, H.M.; Foster, C. Using the SenseCam to Improve Classifications of Sedentary Behavior in Free-Living Settings. *Am. J. Prev. Med.* **2013**, *44*, 290–296.
88. Oliver, M.; Doherty, A.R.; Kelly, P.; Badland, H.M.; Mavoa, S.; Shepherd, J.; Kerr, J.; Marshall, S.; Hamilton, A.; Foster, C. Utility of passive photography to objectively audit built environment features of active transport journeys: An observational study. *Int. J. Health Geogr.* **2013**, *12*, 20.
89. Kelly, P.; Doherty, A.; Mizdrak, A.; Marshall, S.; Kerr, J.; Legge, A.; Godbole, S.; Badland, H.; Oliver, M.; Foster, C. High group level validity but high random error of a self-report travel diary, as assessed by wearable cameras. *J. Transp. Health* **2014**, *1*, 190–201.
90. Barr, M.; Signal, L.; Jenkin, G.; Smith, M. Capturing exposures: Using automated cameras to document environmental determinants of obesity. *Health Promot. Int.* **2014**, *30*, 56–63.
91. Cowburn, G.; Matthews, A.; Doherty, A.; Hamilton, A.; Kelly, P.; Williams, J.; Foster, C.; Nelson, M. Exploring the opportunities for food and drink purchasing and consumption by teenagers during their journeys between home and school: A feasibility study using a novel method. *Public Health Nutr.* **2015**, *19*, 93–103.
92. Signal, L.; Stanley, J.; Smith, M.; Barr, M.; Chambers, T.; Zhou, J.; Duane, A.; Gurrin, C.; Smeaton, A.; McKerchar, C. Children's everyday exposure to food marketing: An objective analysis using wearable cameras. *Int. J. Behav. Nutr. Phys. Act.* **2017**, *14*, 137**.**
93. Liu, W.; Barr, M.; Pearson, A.L.; Chambers, T.; Pfeiffer, K.; Smith, M.; Signal, L. Space-time analysis of unhealthy food advertising: New Zealand children's exposure and health policy options. *Health Promot. Int.* **2019**, *35*, 812–820.
94. McKerchar, C.; Smith, M.; Stanley, J.; Barr, M.; Chambers, T.; Abel, G.; Lacey, C.; Gage, R.; Ni Mhurchu, C.; Signal, L. Food store environment examination–FoodSee: A new method to study the food store environment using wearable cameras. *Glob. Health Promot.* **2019**, *27*, 73–81.
95. Smith, M.; Stanley, J.; Signal, L.; Barr, M.; Chambers, T.; Balina, A.; Mhurchu, C.N.; Wilson, N. Children's healthy and unhealthy beverage availability, purchase and consumption: A wearable camera study. *Appetite* **2019**, *133*, 240–251.
96. Veatupu, L.; Puloka, V.; Smith, M.; McKerchar, C.; Signal, L. Me'akai in Tonga: Exploring the Nature and Context of the Food Tongan Children Eat in Ha'apai Using Wearable Cameras. *Int. J. Environ. Res. Public Health* **2019**, *16*, 1681.
97. Zhou, Q.; Wang, D.; Ni Mhurchu, C.; Gurrin, C.; Zhou, J.; Cheng, Y.; Wang, H. The use of wearable cameras in assessing children's dietary intake and behaviours in China. *Appetite* **2019**, *139*, 1–7.
98. Bulungu, A.L.S.; Palla, L.; Priebe, J.; Forsythe, L.; Katic, P.; Varley, G.; Galinda, B.D.; Sarah, N.; Nambooze, J.; Wellard, K.; et al. Validation of a life-logging wearable camera method and the 24-h diet recall method for assessing maternal and child dietary diversity. *Br. J. Nutr.* **2020**, *125*, 1299–1309.

99. Kohrt, B.; Rai, S.; Vilakazi, K.; Thapa, K.; Bhardwaj, A.; Van Heerden, A. Procedures to Select Digital Sensing Technologies for Passive Data Collection With Children and Their Caregivers: Qualitative Cultural Assessment in South Africa and Nepal. *JMIR Pediatr. Parent.* **2019**, *2*, e12366.
100. Chow, T.E.; Rissman, J. Neurocognitive mechanisms of real-world autobiographical memory retrieval: Insights from studies using wearable camera technology. *Ann. N. Y. Acad. Sci.* **2017**, *1396*, 202–221.
101. Hodges, S.; Williams, L.; Berry, E.; Izadi, S.; Srinivasan, J.; Butler, A.; Smyth, G.; Kapur, N.; Wood, K. SenseCam: A Retrospective Memory Aid. In Proceedings of the International conference on ubiquitous computing Orange County, CA, USA, 17–21 September 2006; pp. 177–193.
102. Kelly, P.; Thomas, E.; Doherty, A.; Harms, T.; Burke, Ó.; Gershuny, J.; Foster, C. Developing a Method to Test the Validity of 24 Hour Time Use Diaries Using Wearable Cameras: A Feasibility Pilot. *PLoS ONE* **2015**, *10*, e0142198.
103. Small, L.; Sidora-Arcoleo, K.; Vaughan, L.; Creed-Capsel, J.; Chung, K.-Y.; Stevens, C. Validity and Reliability of Photographic Diet Diaries for Assessing Dietary Intake Among Young Children. *ICAN Infant Child Adolesc. Nutr.* **2009**, *1*, 27–36.
104. Gershuny, J.; Harms, T.; Doherty, A.; Thomas, E.; Milton, K.; Kelly, P.; Foster, C. *CAPTURE24: Testing Self-Report Time-Use Diaries Against Objective Instruments in Real Time*; Centre for Time Use Research Working Paper: London, UK, 2017.
105. Harms, T.; Gershuny, J.; Doherty, A.; Thomas, E.; Milton, K.; Foster, C. A validation study of the Eurostat harmonised European time use study (HETUS) diary using wearable technology. *BMC Public Health* **2019**, *19*, 455.
106. Bland, J.M.; Altman, D.G. Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement. *Lancet* **1986**, *327*, 307–310.
107. Giavarina, D. Understanding Bland Altman analysis. *Biochem. Med.* **2015**, *25*, 141–151.
108. de Vet, H.C.W.; Terwee, C.B.; Mokkink, L.B.; Knol, D.L. *Measurement in Medicine: A Practical Guide*; Practical Guides to Biostatistics and Epidemiology; Cambridge University Press: Cambridge, NY, USA, 2011; p. 338.
109. Lesotho, B.S. *Household Budget Survey Report, Lesotho*; HBS: Maseru, Lesotho, 2003.
110. Willis, G.B. *Cognitive Interviewing: A "How To" Guide*; Research Triangle Institute: Research Triangle Park, NC, USA, 1999.
111. Ericsson, K.A.; Simon, H.A. Verbal reports as data. *Psychol. Rev.* **1980**, *87*, 215–251.
112. Kelly, P.; Marshall, S.J.; Badland, H.; Kerr, J.; Oliver, M.; Doherty, A.R.; Foster, C. An Ethical Framework for Automated, Wearable Cameras in Health Behavior Research. *Am. J. Prev. Med.* **2013**, *44*, 314–319.
113. Schreiner, M. *Uganda 2012 Poverty Probability Index (PPI)*; Innovations for Poverty Action (IPA): New Haven, CT, USA, 2016.
114. Peterson, R.A. A Meta-Analysis of Cronbach's Coefficient Alpha. *J. Consum. Res.* **1994**, *21*, 381–391.
115. Landis, J.R.; Koch, G.G. The Measurement of Observer Agreement for Categorical Data. *Biometrics* **1977**, *33*, 159.
116. Farooq, M.; Doulah, A.; Parton, J.; McCrory, M.A.; Higgins, J.A.; Sazonov, E. Validation of Sensor-Based Food Intake Detection by Multicamera Video Observation in an Unconstrained Environment. *Nutrients* **2019**, *11*, 609.
117. White, B. Measuring time allocation, decision-making and agrarian changes affecting rural women: Examples from recent research in Indonesia. *IDS Bull.* **1984**, *15*, 18–33.
118. Wilson, G.; Aitken, D.; Hodgson, P.; Bailey, C. The hidden impact of home adaptations: Using a wearable camera to explore lived experiences and taken-for-granted behaviours. *Health Soc. Care Community* **2019**, *27*, 1469–1480.
119. Floro, M.S.; Miles, M. Time use, work and overlapping activities: Evidence from Australia. *Camb. J. Econ.* **2003**, *27*, 881–904.
120. Rost, L.; Bates, K.; Dellepiane, L. *Women's Economic Empowerment and Care: Evidence for Influencing*; Oxfam GB: Oxford, UK, 2015.
121. Mueller, C. *Time Use Data: Sources and Applications of Data on Paid and Unpaid Labour*; International Labour Organization: Geneva, Switzerland, 2018.
122. Lentz, E.; Kerr, R.B.; Patel, R.; Dakishoni, L.; Lupafya, E. The Invisible Hand that Rocks the Cradle: On the Limits of Time Use Surveys. *Dev. Chang.* **2018**, *50*, 301–328.
123. Kan, M.Y.; Pudney, S. Measurement Error in Stylized and Diary Data on Time Use. *Sociol. Methodol.* **2008**, *38*, 101–132.
124. Raber, M.; Patterson, M.; Jia, W.; Sun, M.; Baranowski, T. Utility of eButton images for identifying food preparation behaviors and meal-related tasks in adolescents. *Nutr. J.* **2018**, *17*, 32.
125. Lee, R.; Skinner, A.; Bornstein, M.H.; Radford, A.N.; Campbell, A.; Graham, K.; Pearson, R.M. Through babies' eyes: Practical and theoretical considerations of using wearable technology to measure parent–infant behaviour from the mothers' and infants' view points. *Infant. Behav. Dev.* **2017**, *47*, 62–71.
126. Raber, M.; Baranowski, T.; Crawford, K.; Sharma, S.V.; Schick, V.; Markham, C.; Jia, W.; Sun, M.; Steinman, E.; Chandra, J. The Healthy Cooking Index: Nutrition Optimizing Home Food Preparation Practices across Multiple Data Collection Methods. *J. Acad. Nutr. Diet.* **2020**, *120*, 1119–1132.
127. Tucker, K.; Sanjur, D. Maternal employment and child nutrition in Panama. *Soc. Sci. Med.* **1988**, *26*, 605–612.
128. Wandel, M.; Holmboe-Ottesen, G. Women's Work in Agriculture and Child Nutrition in Tanzania. *J. Trop. Pediatr.* **1992**, *38*, 252–255.
129. Laskaris, Z.; Milando, C.; Batterman, S.; Mukherjee, B.; Basu, N.; O'Neill, M.S.; Robins, T.G.; Fobil, J.N. Derivation of Time-Activity Data Using Wearable Cameras and Measures of Personal Inhalation Exposure among Workers at an Informal Electronic-Waste Recovery Site in Ghana. *Ann. Work Expo. Health* **2019**, *63*, 829–841.
130. Alharbi, R.; Stump, T.; Vafaie, N.; Pfammatter, A.; Spring, B.; Alshurafa, N. I Can't Be Myself. *Proc ACM Interact. Mob. Wearable Ubiquitous Technol.* **2018**, *2*, 1–40.

131. Chambers, T.; Pearson, A.; Kawachi, I.; Rzotkiewicz, Z.; Stanley, J.; Smith, M.; Barr, M.; Ni Mhurchu, C.; Signal, L. Kids in space: Measuring children's residential neighborhoods and other destinations using activity space GPS and wearable camera data. *Soc. Sci. Med.* **2017**, *193*, 41–50.
132. Gelonch, O.; Ribera, M.; Codern-Bové, N.; Ramos, S.; Quintana, M.; Chico, G.; Cerulla, N.; Lafarga, P.; Radeva, P.; Garolera, M. Acceptability of a lifelogging wearable camera in older adults with mild cognitive impairment: A mixed-method study. *BMC Geriatr.* **2019**, *19*, 110.
133. Sellen, A.J.; Fogg, A.; Aitken, M.; Hodges, S.; Rother, C.; Wood, K. Do Life-Logging Technologies Support Memory for the Past? In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA, 30 April–6 May 2007; pp. 81–90.
134. Wilson, G.; Jones, D.; Schofield, P.; Martin, D.J. Experiences of using a wearable camera to record activity, participation and health-related behaviours: Qualitative reflections of using the Sensecam. *Digit. Health* **2016**, *2*, 205520761668262.
135. Grosh, M.; Glewwe, P. *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study*; Report No.: 20731; The World Bank: Washington, DC, USA, 2000; Volume 3.
136. Malapit, H.; Quisumbing, A.; Meinzen-Dick, R.; Seymour, G.; Martinez, E.M.; Heckert, J.; Rubin, D.; Vaz, A.; Yount, K.M. Development of the project-level Women's Empowerment in Agriculture Index (pro-WEAI). *World Dev.* **2019**, *122*, 675–692.
137. International Food Policy Research Institute. *Project-Level Women's Empowerment in Agriculture Index (Pro-WEAI) Pilot Version*; International Food Policy Research Institute: Washington, DC, USA, 2020.

Supplementary Materials

Table S1. Household data collection patterns.

Pattern	DAY 1	DAY 2	DAY 3	DAY 4	DAY 5
1	ICF MQ Anthro	OBS AWC	24HR IAR	AWC	IAR TQ
2	ICF MQ Anthro	AWC	IAR	OBS AWC	24HR IAR TQ

ICF, informed consent form; MQ, mothers' questionnaire; Anthro, anthropometry; OBS, observation; AWC, automated wearable camera; 24HR, 24-hour recall; IAR, image-assisted recall; TQ, technology questionnaire.

Table S2. Time use activities and ICATUS major divisions.

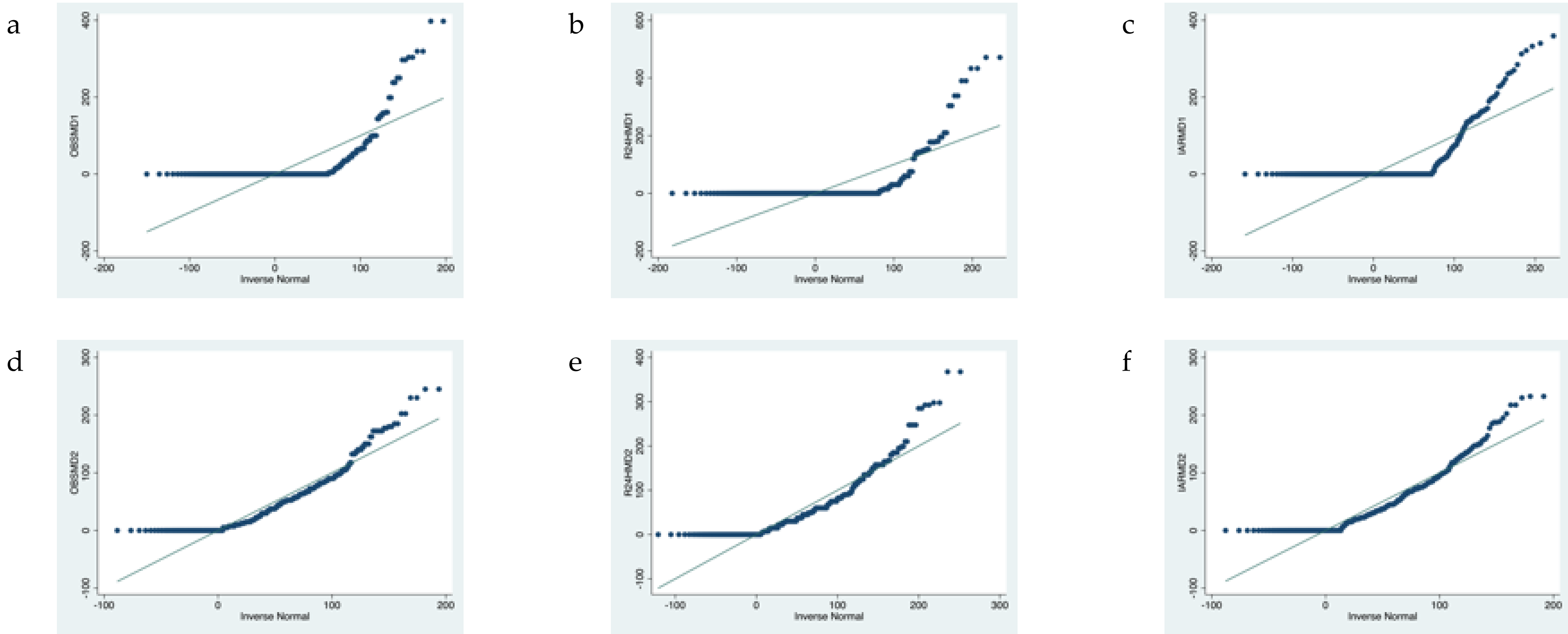
ICATUS-2016 Major Divisions*		1 Employment and related activities	2 Production of goods for own final use	3 Unpaid domestic services for household and family members	4 Unpaid caregiving services for household and family members	5 Unpaid volunteer, trainee, and other unpaid work †	6 Learning	7 Socializing and communication, community participation	8 Culture, leisure, mass media and sports practices	9 Self-care and maintenance
1	Cash crop farming	X								
2	Cooking food - business	X								
3	Working – employed	X								
4	Working - own business	X								
5	Supervising employees	X								
6	Food crop farming		X							
7	Fish farming		X							
8	Livestock rearing		X							
9	Fish-livestock product processing		X							
10	Shelling maize – hand		X							
11	Shelling maize –machine		X							
12	Wild food gathering & fishing		X							
13	Fetching fuel (incl. travel)		X							
14	Fetching water (incl. travel)		X							
15	Chopping / splitting firewood			X						

ICATUS-2016 Major Divisions*		1 Employment and related activities	2 Production of goods for own final use	3 Unpaid domestic services for household and family members	4 Unpaid caregiving services for household and family members	5 Unpaid volunteer, trainee, and other unpaid work †	6 Learning	7 Socializing and communication, community participation	8 Culture, leisure, mass media and sports practices	9 Self-care and maintenance
16	Other domestic work			X						
17	Other shopping			X						
18	Washing clothes			X						
19	Cooking food - family / friends			X						
20	Food collection – garden / farm			X						
21	Food collection – shop / vendor			X						
22	Food preparation			X						
23	Other post-harvest processing			X						
24	Travel to / from farm - food			X						
25	Travel to / from market - food			X						
26	Other travel			X						
27	Care of index‡				X					
28	Care of other children or adults‡				X					
29	Accessing services				X					
30	Breast feeding index				X					
31	Feeding index food or drinks				X					

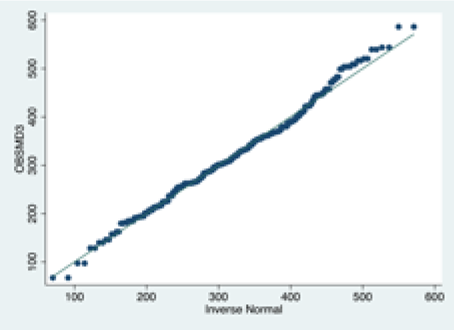
ICATUS-2016 Major Divisions*		1 Employment and related activities	2 Production of goods for own final use	3 Unpaid domestic services for household and family members	4 Unpaid caregiving services for household and family members	5 Unpaid volunteer, trainee, and other unpaid work †	6 Learning	7 Socializing and communication, community participation	8 Culture, leisure, mass media and sports practices	9 Self-care and maintenance
32	Feeding other children or adults				X					
33	Playing w index				X					
34	Studying						X			
35	Chatting w friends / relatives‡							X		
36	Religious activities							X		
37	Social activities & hobbies							X		
38	TV/radio/reading‡								X	
39	Personal care									X
40	Eating or drinking									X
41	Resting§									
42	Sleeping§									
43	Answering the phone for the study§									
44	Other interactions for the study§									

* United Nations Statistics Division. (2019). International Classification of Activities for Time-Use Statistics 2016 (ICATUS 2016). New York: United Nations.
† There were no activities that mapped to ICATUS MD5, Unpaid volunteer, trainee, and other unpaid work.
‡ Simultaneous activity
§ The following activities were not included in the analysis: resting or sleeping (due to known inconsistencies with recording) or answering the phone for the study or other interactions for the study.

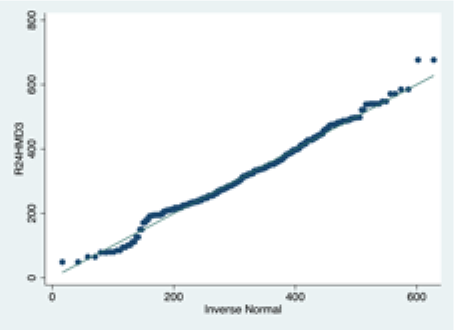
Figure S1. Plots of the quantiles of activities against the quantiles of the normal distribution (Q-Q plot) for employment and related activities [MD1] (a) OBS, (b) 24HR, and (c) IAR; production of goods for own final use [MD2]] (c) OBS, (e) 24HR, and (f) IAR; unpaid domestic services for household and family members [MD3]] (g) OBS, (h) 24HR, and (i) IAR; unpaid caregiving services for household and family members [MD4]] (j) OBS, (k) 24HR, and (l) IAR; socializing and communication, community participation and religious practice [MD7]] (m) OBS, (n) 24HR, and (o) IAR; culture, leisure, mass media and sports practice [MD8]] (p) OBS, (q) 24HR, and (r) IAR; and self-care and maintenance [MD9]] (s) OBS, (t) 24HR, and (u) IAR.



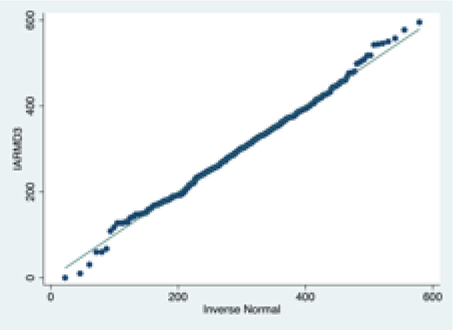
g


OBSMD3
Inverse Normal

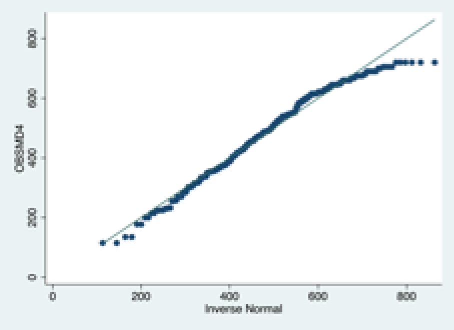
h


R24HMD3
Inverse Normal

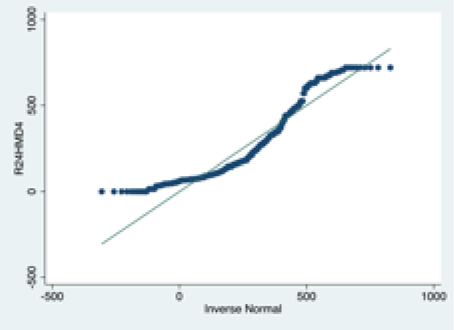
i


IARMD3
Inverse Normal

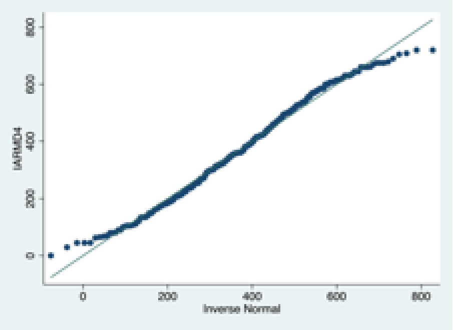
j


OBSMD4
Inverse Normal

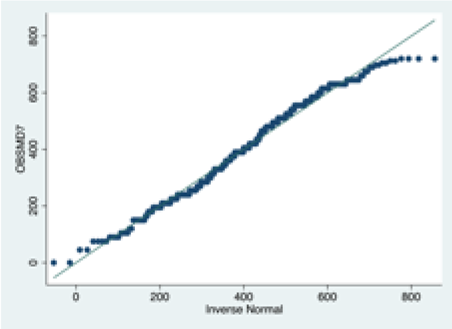
k


R24HMD4
Inverse Normal

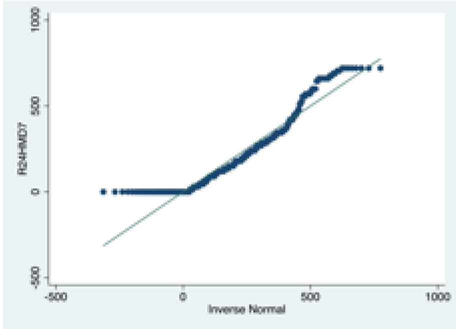
l


IARMD4
Inverse Normal

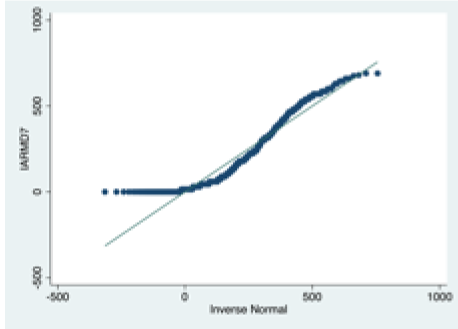
m


OBSMD7
0
200
400
600
800
Inverse Normal

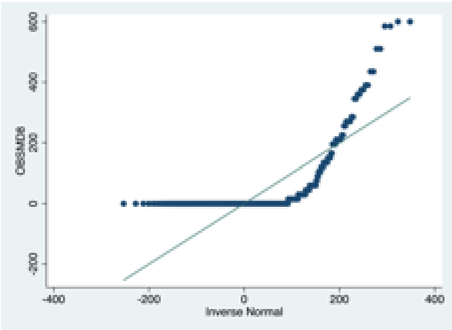
n


R24HMD7
-500
0
500
1000
Inverse Normal

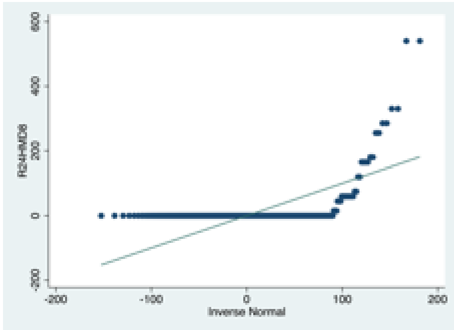
o


IARMD7
-500
0
500
1000
Inverse Normal

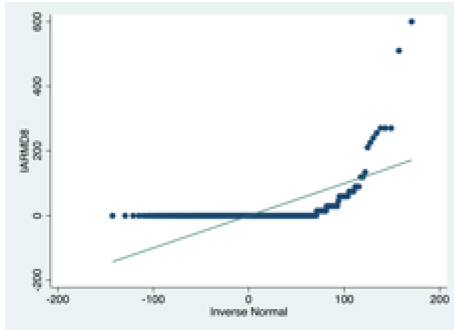
p


OBSMD8
-200
0
200
400
600
-400
Inverse Normal

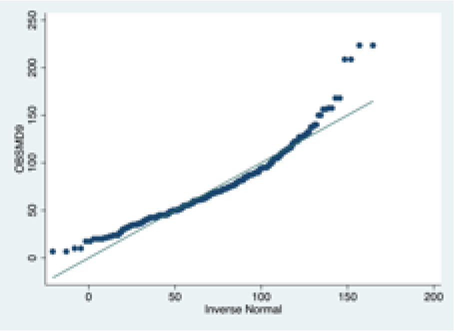
q


R24HMD8
-200
0
200
400
600
-100
100
Inverse Normal

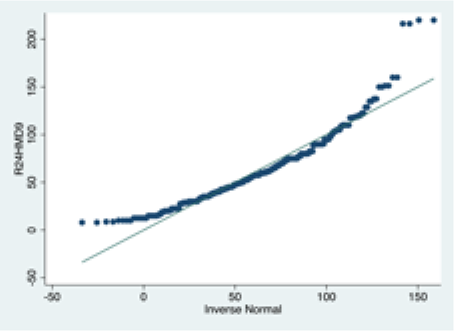
r


IARMD8
-200
0
200
400
600
-100
100
Inverse Normal

s


OBSMD9
Inverse Normal

t


R24HMD9
Inverse Normal

u

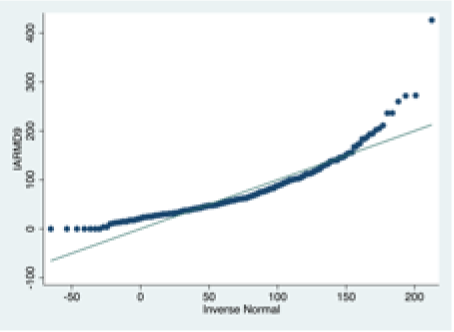

IARMD9
Inverse Normal

Table S3. Inter-method comparison of the median time allocated in minutes to activity groups, including enumerator image interpretation (EII).

ICATUS Activity Group	OBS (N=175)		24HR (N=175)				EII (N=175)				IAR (N=175)			
	Median (minutes)	25th, 75th	Median (minutes)	25th, 75th	MD	P	Median (minutes)	25th, 75th	MD	P	Median (minutes)	25th, 75th	MD	P
Employment and related activities [MD1]	0	0, 5	0	0, 0	0	0.7987	0	0, 0	0	0.6050	0	0, 35	0	0.1123
Production of goods for own final use [MD2]	45	10, 79	49	15, 90	-3	0.0043	35	0, 75	5	0.0230	43	18, 81	0	0.8619
Unpaid domestic services for household and family members [MD3]	318	263, 370	320	245, 396	6	0.7903	295	218, 355	25	<0.0001	311	251, 374	8	0.1836
Unpaid caregiving services for household and family member [MD4] *	491	388, 608	180	96, 390	264	<0.0001	315	208, 445	153	<0.0001	418	324, 541	56	<0.0001
Socializing and communication, community participation and religious practice [MD7] *	405	270, 525	195	75, 330	180	<0.0001	90	15, 225	225	<0.0001	285	105, 465	105	<0.0001
Culture, leisure, mass media and sports practices [MD8] *	0	0, 30	0	0, 0	0	<0.0001	0	0, 0	0	<0.0001	0	0, 0	0	<0.0001
Self-care and maintenance [MD9]	68	50, 88	58	39, 80	11	0.0010	49	31, 78	14	<0.0001	79	53, 111	-9	0.0002

OBS, observation; 24HR, 24-hour recall; EII, enumerator image interpretation; IAR, image-assisted recall; MD, median of the time allocation differences (versus OBS); P, p-value of Wilcoxon signed rank sum test of time allocated compared to OBS.
* Activity group contains one or more simultaneous activities.

Table S4. Inter-method comparison of the median time allocated in minutes to activity groups for participating women only.

ICATUS Activity Group	N	OBS		24HR		IAR	
		Median (minutes)	25th, 75th	Median (minutes)	25th, 75th	Median (minutes)	25th, 75th
Employment and related activities [MD1]	77	16	0, 64	10	0, 60	53	4, 135
Production of goods for own final use [MD2]	159	50	15, 83	60	26, 94	45	24, 85
Culture, leisure, mass media and sports practices [MD8] *	73	45	15, 165	0	0, 0	0	0, 60

OBS, observation; 24HR, 24-hour recall; IAR, image-assisted recall.

* Activity group contains one or more simultaneous activities.

NB: There were no activities that mapped to ICATUS MD5, Unpaid volunteer, training, or other unpaid work.

NB: This table does not include resting or sleeping (due to known inconsistencies with recording) or answering the phone for the study or other interactions for the study.

Table S5. Inter-method comparison of the median number of concurrent activities.

	N	OBS		24HR		IAR	
		Median	25th, 75th	Median	25th, 75th	Median	25th, 75th
Median concurrent activities	175	3.0	2.5, 4.0	2.0†	1.0, 2.0	3.0	2.0, 4.0

OBS, observation; 24HR, 24-hour recall; IAR, image-assisted recall.

† P-value of Wilcoxon signed rank sum test of median time allocated compared to OBS <0.05.

NB: This data reflects the medians across the 40 discrete activities, not (the 9) activity groups.

NB: This table does NOT include resting or sleeping (due to known inconsistencies with recording) or answering the phone for the study or other interactions for the study.

Table S6. Inter-method comparison of the median number and proportion of timeslots containing concurrent activities.

Timeslots	OBS	24HR	IAR
Median number (out of 48)	42	26	40
Median proportion	88%	54%	83%

OBS, observation; 24HR, 24-hour recall; IAR, image-assisted recall.

Table S7a. Inter-method comparison of the median time allocated in minutes to activity groups, observation day verses non-observation day.

ICATUS Activity Group	NON OBS EII (N=174)		OBS EII (N=175)		P	NON OBS IAR (N=174)		OBS IAR (N=175)		P
	Median	25th, 75th	Median	25th, 75th		Median	25th, 75th	Median	25th, 75th	
Employment and related activities [MD1]	0	0, 8	0	0, 0	0.0691	0	0, 26	0	0, 35	0.6984
Production of goods for own final use [MD2]	53	8, 104	35	0, 75	0.0005	60	16, 109	43	18, 81	0.0021
Unpaid domestic services for household and family members [MD3]	267	188, 345	295	218, 355	0.0525	296	224, 376	311	251, 374	0.0678
Unpaid caregiving services for household and family member [MD4] *	235	140, 398	315	208, 445	0.0014	339	206, 480	418	324, 541	<0.0001
Socializing and communication, community participation & religious practice [MD7] *	150	45, 285	90	15, 225	0.0141	278	150, 450	285	105, 465	0.3588
Culture, leisure, mass media and sports practices [MD8] *	0	0, 0	0	0, 0	0.6070	0	0, 0	0	0, 0	0.4107
Self-care and maintenance [MD9]	49	28, 79	49	31, 76	0.4350	74	44, 110	79	53, 111	0.2469

NON-OBS EII, non-observation day enumerator image interpretation; OBS EII, observation day enumerator image interpretation; P, p-value of Wilcoxon signed rank sum test of time allocated, non-observation day compared to observation day; NON-OBS IAR, non-observation day image-assisted recall; OBS IAR, observation day image-assisted recall.
* Activity group contains one or more simultaneous activities.

Table S8. Frequency of inter-method time allocation differences.

ICATUS Activity Group	Method	Frequency of no differences N (%)	Frequency of differences <15 min N (%)	Frequency of differences <30 min N (%)	Frequency of differences >1 hr N (%)	Frequency of differences >2 hr N (%)
Employment and related activities [MD1]	24HR	114 (65)	126 (72)	138 (79)	26 (15)	11 (6)
	IAR	105 (60)	118 (67)	132 (75)	23 (13)	8 (5)
Production of goods for own final use [MD2]	24HR	26 (15)	74 (42)	111 (63)	27 (15)	6 (3)
	IAR	29 (17)	84 (48)	121 (69)	17 (10)	5 (3)
Unpaid domestic services for household and family members [MD3]	24HR	1 (1)	20 (11)	43 (25)	98 (56)	46 (26)
	IAR	1 (1)	29 (17)	54 (31)	73 (42)	19 (11)
Unpaid caregiving services for household and family member [MD4] *	24HR	0 (0)	7 (4)	9 (5)	159 (91)	139 (79)
	IAR	1 (1)	10 (6)	25 (14)	125 (71)	84 (48)
Socializing and communication, community participation and religious practice [MD7] *	24HR	5 (3)	5 (3)	10 (6)	151 (86)	134 (77)
	IAR	4 (2)	4 (2)	12 (7)	144 (82)	112 (64)
Culture, leisure, mass media and sports practices [MD8] *	24HR	109 (62)	109 (62)	123 (70)	37 (21)	25 (14)
	IAR	110 (63)	110 (63)	125 (71)	36 (21)	22 (13)
Self-care and maintenance [MD9]	24HR	1 (1)	52 (30)	97 (55)	19 (11)	4 (2)
	IAR	2 (1)	46 (26)	96 (55)	30 (17)	11 (6)

min, minute; hr, hour.

* Activity group contains one or more simultaneous activities.

Table S9. Frequency of inter-method median concurrent activities differences.

	Method	Frequency of no differences N (%)	Frequency of differences <1 CA N (%)	Frequency of differences <2 CA N (%)	Frequency of differences >4 CA N (%)
Median concurrent activities	24HR	30 (17)	43 (25)	92 (53)	1 (1)
	IAR	37 (21)	55 (31)	129 (74)	8 (5)

CA, concurrent activities; 24HR, 24-hour recall; IAR, image-assisted recall.

Table S10. Inter-method comparison of the median concurrent activities bias and limits of agreement (LOA).

	Bias* (activities)	LOA†	
24HR	1.3	-1.1	3.7
IAR	0.0	-3.2	3.2

LOA, limits of agreement; 24HR, 24-hour recall; IAR, image-assisted recall.
* Mean difference.
† +/- 2 SD from the mean difference.

NB: A negative indicates that 24R / IAR overestimated OBS.

Table S11. Inter-method comparison of reliability for median concurrent activities.

	24HR					IAR				
	% Agree.	Cohen's kappa†	P	95% CI		% Agree.	Cohen's kappa†	P	95% CI	
Median concurrent activities	0.7091	0.0281	0.198	-0.0149	0.0710	0.8521	0.0308	0.487	-0.0565	0.1181

24HR, 24-hour recall; IAR, image-assisted recall; CI, confidence interval.
† Using weighted Cohen's (reliability) kappa. Landis and Koch (1977) suggest the following benchmark scale for interpreting the kappa statistic: <0.00 Poor; 0.00-0.20 Slight; 0.21-0.40 Fair; 0.41-0.60 Moderate; 0.61-0.80 Substantial; 0.81-1.00 Almost Perfect.

Table S12. Inter-method comparison of the median time allocated in minutes to discrete activities.

		OBS		24HR				EII				IAR			
	N	*Median*	*25th, 75th*	*Median*	*25th, 75th*	*MD*	*P*	*Median*	*25th, 75th*	*MD*	*P*	*Median*	*25th, 75th*	*MD*	*P*
ICATUS MD1 - Employment and related activities																
Cash crop farming	175	0	0, 0	0	0, 0	0	0.7550	0	0, 0	0	0.0334	0	0, 0	0	0.2040
Cooking food - business	175	0	0, 0	0	0, 0	0	0.2328	0	0, 0	0	0.0009	0	0, 0	0	0.0067
Working – employed	175	0	0, 0	0	0, 0	0	0.9968	0	0, 0	0	0.1573	0	0, 0	0	0.1573
Working - own business	175	0	0, 0	0	0, 0	0	0.8234	0	0, 0	0	0.0616	0	0, 35	0	0.0003
Supervising employees	175	0	0, 0	0	0, 0	0	0.0833	0	0, 0	0	0.0833	0	0, 0	0	0.0833
ICATUS MD2 - Production of goods for own final use																
Food crop farming	175	0	0, 0	0	0, 0	0	0.1136	0	0, 15	0	0.0740	0	0, 0	0	0.5500
Fish farming	175	0	0, 0	0	0, 0	0	0.3173	0	0, 0	0	.	0	0, 0	0	.
Livestock rearing	175	0	0, 5	0	0, 0	0	0.4705	0	0, 0	0	<0.0001	0	0, 8	0	0.9048
Fish-livestock product processing	175	0	0, 0	0	0, 0	0	0.6491	0	0, 0	0	0.1573	0	0, 0	0	0.9954
Shelling maize – hand	175	0	0, 0	0	0, 0	0	0.9968	0	0, 0	0	0.1573	0	0, 0	0	0.1573
Shelling maize –machine	175	0	0, 0	0	0, 0	0	0.1831	0	0, 0	0	0.0455	0	0, 0	0	0.1764
Wild food gathering & fishing	175	0	0, 0	0	0, 0	0	0.3173	0	0, 0	0	0.3173	0	0, 0	0	0.3173
Fetching fuel (incl. travel)	175	0	0, 8	0	0, 0	0	0.4273	0	0, 0	0	0.0003	0	0, 10	0	0.8883
Fetching water (incl. travel)	175	20	0, 45	28	0, 60	0	0.0069	13	0, 43	0	0.2054	20	0, 45	0	0.8087
ICATUS MD3 - Unpaid domestic services for household and family members																
Chopping / splitting firewood	175	0	0, 8	0	0, 0	0	<0.0001	0	0, 0	0	<0.0001	0	0, 0	0	0.0014
Other domestic work	175	83	50, 114	53	23, 90	27	<0.0001	58	25, 83	23	<0.0001	68	34, 98	19	<0.0001
Other shopping	175	0	0, 0	0	0, 0	0	0.1508	0	0, 0	0	0.0013	0	0, 0	0	0.0137
Washing clothes	175	0	0, 20	0	0, 15	0	0.9219	0	0, 23	0	0.0165	0	0, 26	0	0.6586
Cooking food - family / friends	175	85	56, 111	100	60, 158	-15	0.0001	64	38, 96	20	0.0001	70	48, 100	13	0.0023
Food collection – garden / farm	175	0	0, 15	0	0, 15	0	0.2274	0	0, 0	0	0.0002	0	0, 15	0	0.9843
Food collection – shop / vendor	175	0	0, 0	0	0, 0	0	0.6527	0	0, 0	0	0.0013	0	0, 5	0	0.0755
Food preparation	175	51	30, 74	45	15, 73	4	0.1814	48	23, 78	0	0.9477	49	26, 79	1	0.9976
Other post-harvest processing	175	0	0, 13	0	0, 0	0	<0.0001	0	0, 24	0	0.0023	0	0, 28	0	<0.0001

	OBS			24HR				EII				IAR			
	N	*Median*	*25th, 75th*	*Median*	*25th, 75th*	*MD*	*P*	*Median*	*25th, 75th*	*MD*	*P*	*Median*	*25th, 75th*	*MD*	*P*
Travel to / from farm - food	175	0	0, 20	0	0, 15	0	0.6650	0	0, 0	0	0.0056	0	0, 13	0	0.0352
Travel to / from market - food	175	0	0, 0	0	0, 0	0	0.7101	0	0, 0	0	0.5086	0	0, 11	0	0.0003
Other travel	175	11	0, 33	0	0, 23	0	0.0259	15	0, 53	0	0.0017	15	0, 53	0	0.0043
ICATUS MD4 - Unpaid caregiving services for household and family members															
Care of index *	175	405	285, 510	105	30, 300	225	<0.0001	240	135, 360	120	<0.0001	315	210, 465	60	<0.0001
Care of other children or adults *	175	255	105, 390	30	0, 150	150	<0.0001	120	45, 255	90	<0.0001	165	45, 270	45	<0.0001
Accessing services	175	0	0, 0	0	0, 0	0	0.6585	0	0, 0	0	0.0254	0	0, 0	0	0.9943
Breast feeding index	175	44	0, 78	0	0, 23	15	<0.0001	0	0, 0	26	<0.0001	0	0, 45	0	<0.0001
Feeding index food or drinks	175	35	21, 57	20	10, 33	14	<0.0001	31	15, 50	4	0.1069	46	28, 70	-9	0.0001
Feeding other children or adults	175	21	10, 33	15	5, 26	6	0.0001	19	8, 33	0	0.6409	25	11, 38	-2	0.1400
Playing w index	175	15	0, 35	0	0, 0	13	<0.0001	0	0, 0	13	<0.0001	0	0, 15	8	<0.0001
ICATUS MD6 - Learning															
Studying	175	0	0, 0	0	0, 0	0	.	0	0, 0	0	.	0	0, 0	0	.
ICATUS MD7 - Socializing and communication, community participation and religious practice															
Chatting w friends / relatives *	175	390	270, 525	180	75, 315	195	<0.0001	90	15, 225	225	<0.0001	270	105, 450	105	<0.0001
Religious activities	175	0	0, 0	0	0, 0	0	0.2514	0	0, 0	0	0.0254	0	0, 0	0	0.7202
Social activities & hobbies	175	0	0, 0	0	0, 0	0	0.2302	0	0, 0	0	0.0044	0	0, 0	0	0.5385
ICATUS MD8 - Culture, leisure, mass media and sports practices															
TV/radio/reading *	175	0	0, 30	0	0, 0	0	<0.0001	0	0, 0	0	<0.0001	0	0, 0	0	<0.0001
ICATUS MD9 - Self-care and maintenance															
Personal care	175	30	19, 45	26	8, 38	8	0.0024	0	0, 13	23	<0.0001	29	15, 49	1	0.4473
Eating or drinking	175	35	24, 48	30	19, 45	3	0.0819	38	24, 63	-6	0.0071	40	28, 63	-6	<0.0001

OBS, observation; 24HR, 24-hour recall; EII, enumerator image interpretation; IAR, image-assisted recall; MD, mean proportion difference (versus OBS); P, p-value of Wilcoxon signed rank sum compared to OBS.
* Simultaneous activity.
NB: There were no activities that mapped to ICATUS MD5, Unpaid volunteer, trainee, and other unpaid work.
NB: This table does not include resting or sleeping (due to known inconsistencies with recording) or answering the phone for the study or other interactions for the study.

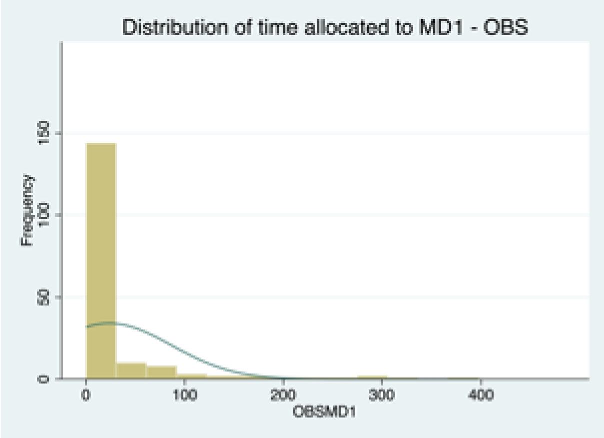
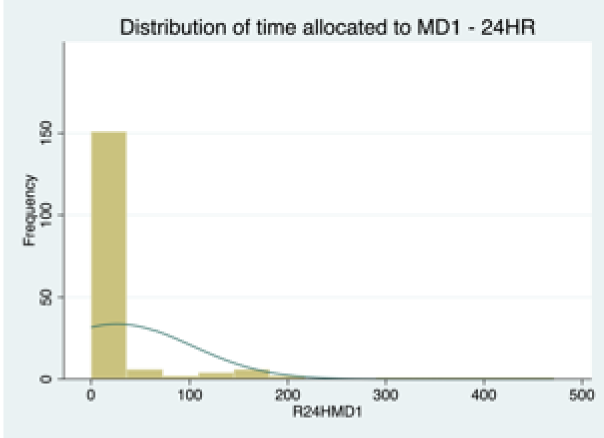
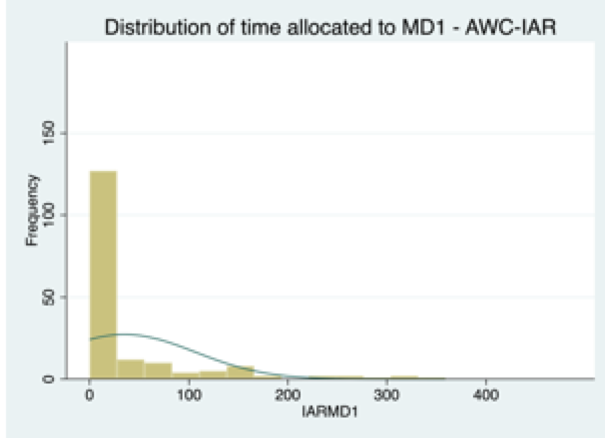
Table S13. Within-method comparison of time allocation - with and without simultaneous activities.

	Allowing for SA		Traditional Approach		% Change	Allowing for SA		Traditional Approach		% Change	Allowing for SA		Traditional Approach		% Change
ICATUS Activity Group	OBS		OBS			24HR		24HR			IAR		IAR		
	Median (minutes)	25th, 75th	Median (minutes)	25th, 75th		Median (minutes)	25th, 75th	Median (minutes)	25th, 75th		Median (minutes)	25th, 75th	Median (minutes)	25th, 75th	
Employment and related activities [MD1]	0	0, 5	0	0, 3	0%	0	0, 0	0	0, 0	0%	0	0, 35	0	0, 24	0%
Production of goods for own final use [MD2]	45	10, 79	26	6, 43	-42%	49	15, 90	38	8, 75	-22%	43	18, 81	29	11, 61	-33%
Unpaid domestic services for household and family members [MD3]	318	263, 370	185	143, 223	-42%	320	245, 396	250	180, 318	-23%	311	251, 374	208	168, 275	-33%
Unpaid caregiving services for household and family member [MD4]	491	388, 608	274	211, 341	-44%	180	96, 390	140	75, 223	-22%	418	324, 541	237	194, 299	43%
Socializing and communication, community participation and religious practice [MD7]	405	270, 525	137	89, 165	-66%	195	75, 330	108	38, 169	-47%	285	105, 465	97	39, 140	-66%
Culture, leisure, mass media and sports practices [MD8]	0	0, 30	0	0, 8	0%	0	0, 0	0	0, 0	0%	0	0, 0	0	0, 0	0%
Self-care and maintenance [MD9]	68	50, 88	41	31, 57	-40%	58	39, 80	48	30, 68	-17%	79	53, 111	57	40, 82	-28%

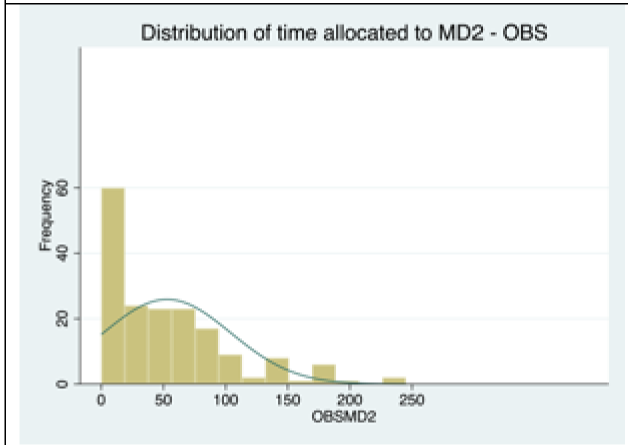
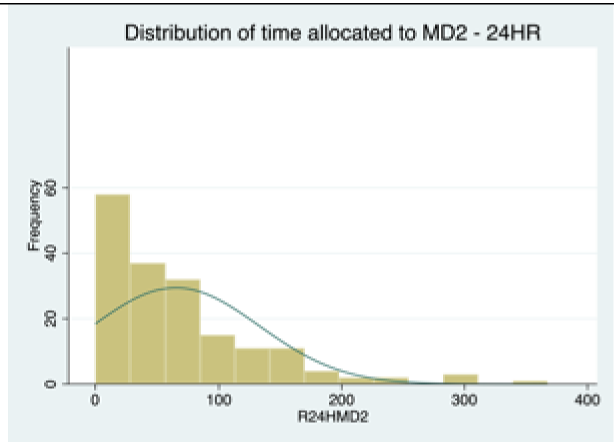
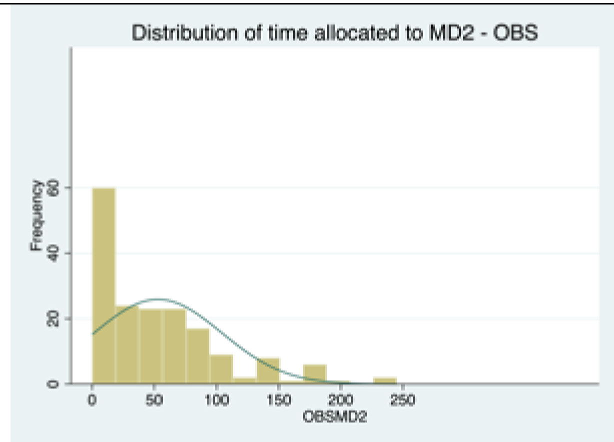
OBS, observation; 24HR, 24-hour recall; EII, enumerator image interpretation; IAR, image-assisted recall; MD, mean proportion difference (versus OBS); P, p-value of Wilcoxon signed rank sum compared to OBS.

ADDITIONAL MATERIALS (NOT INCLUDED IN THE MANUSCRIPT)

Additional Figure 1. Inter-method comparison of the distribution of time allocated to employment and related activities.

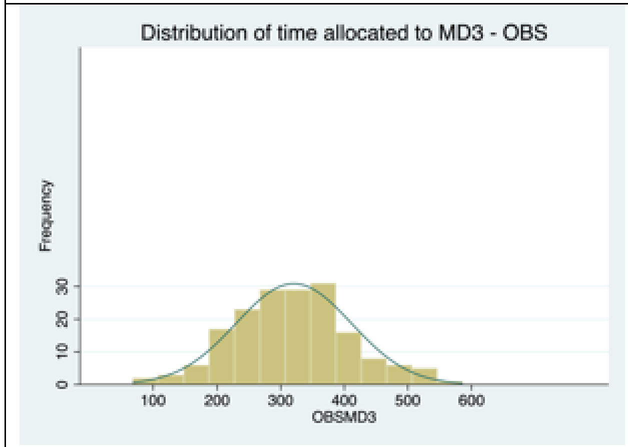
OBS	24HR	AWC-IAR
		
OBS, observation; 24HR, 24-hour recall; AWC-IAR, automated wearable camera-based image-assisted recall.		

Additional Figure 2. Inter-method comparison of the distribution of time allocated to Production of goods for own final use.

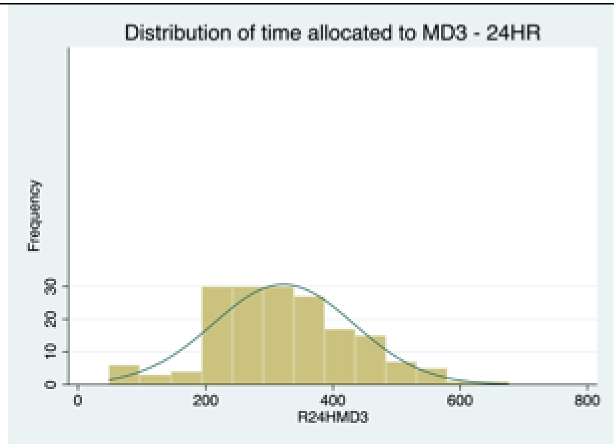
OBS	24HR	AWC-IAR
		
OBS, observation; 24HR, 24-hour recall; AWC-IAR, automated wearable camera-based image-assisted recall.		

Additional Figure 3. Inter-method comparison of the distribution of time allocated to unpaid domestic services for household and family members.

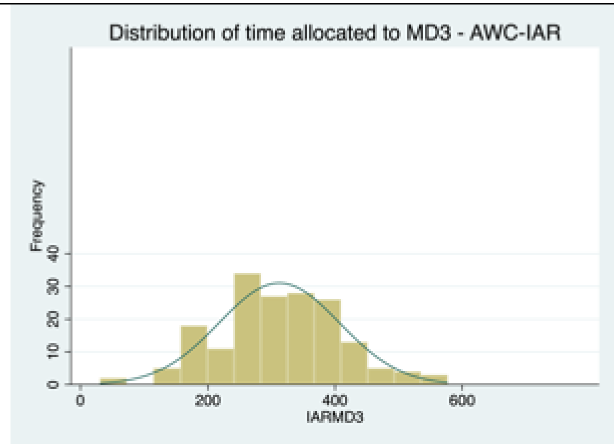
OBS



24HR



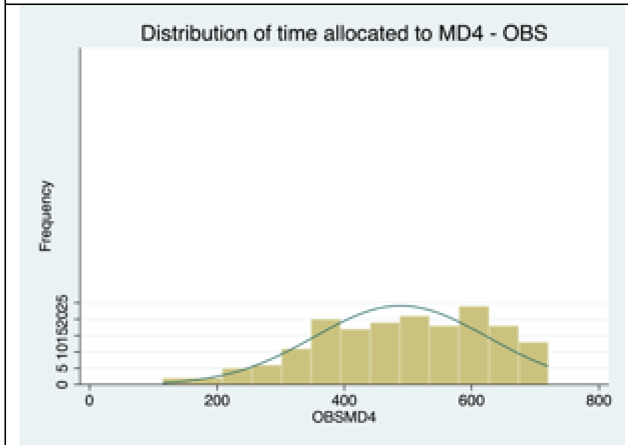
AWC-IAR



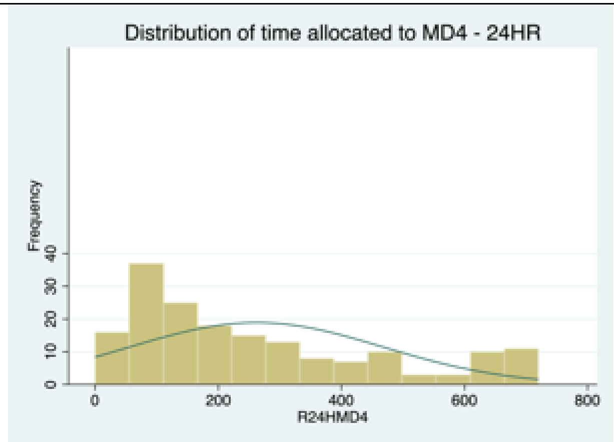
OBS, observation; 24HR, 24-hour recall; AWC-IAR, automated wearable camera-based image-assisted recall.

Additional Figure 4. Inter-method comparison of the distribution of time allocated to unpaid caregiving services for household and family members.

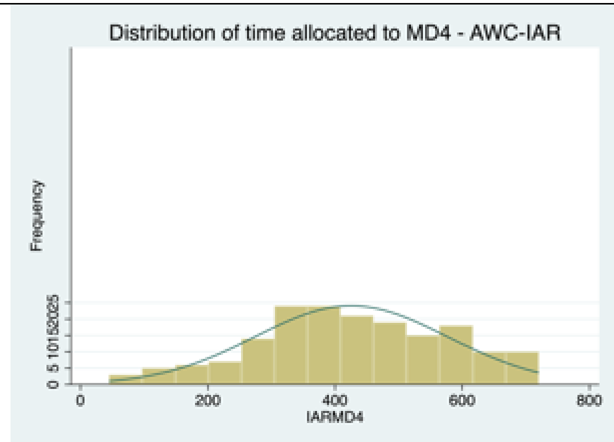
OBS



24HR

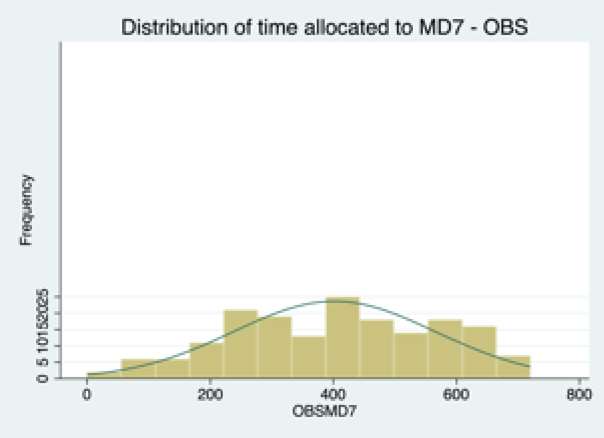
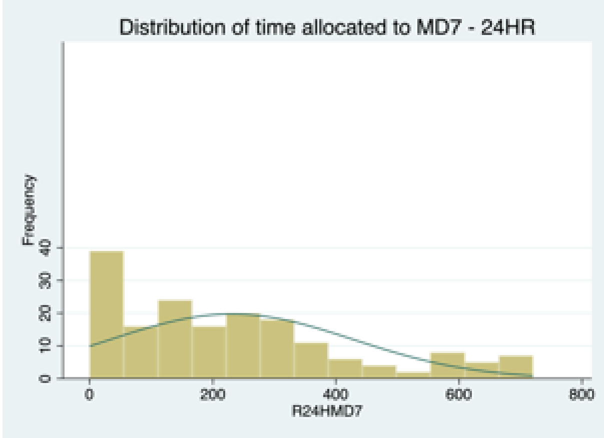
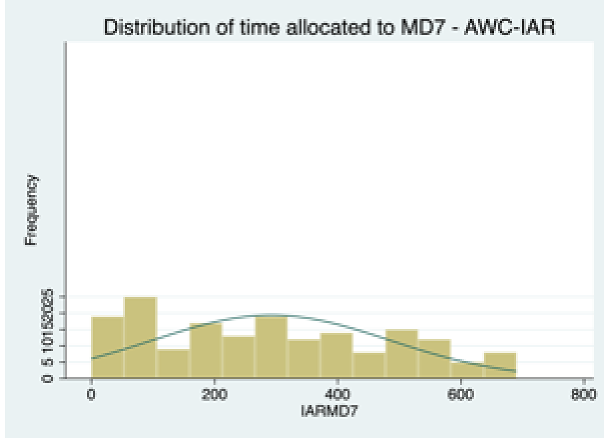


AWC-IAR



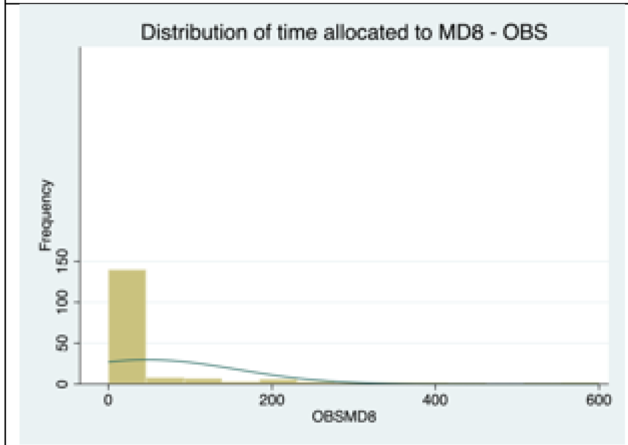
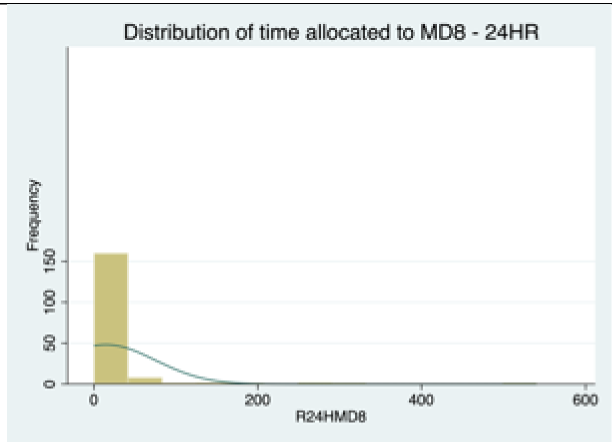
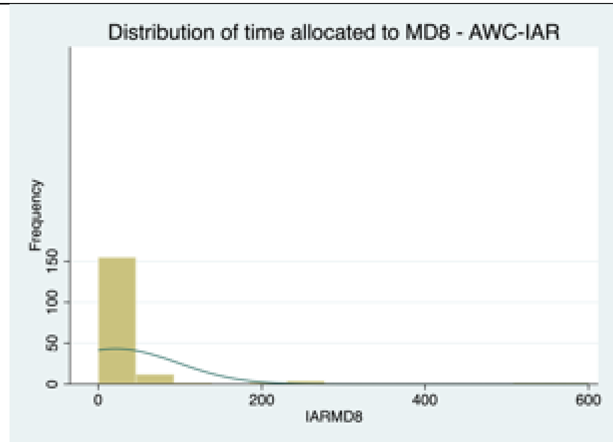
OBS, observation; 24HR, 24-hour recall; AWC-IAR, automated wearable camera-based image-assisted recall.

Additional Figure 5. Inter-method comparison of the distribution of time allocated to socializing and communication, community participation and religious practice.

OBS	24HR	AWC-IAR
		

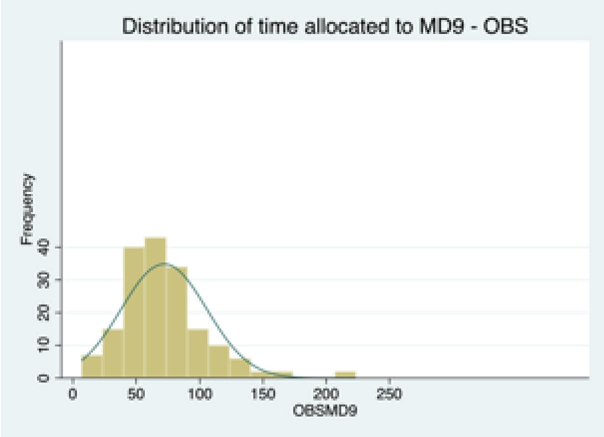
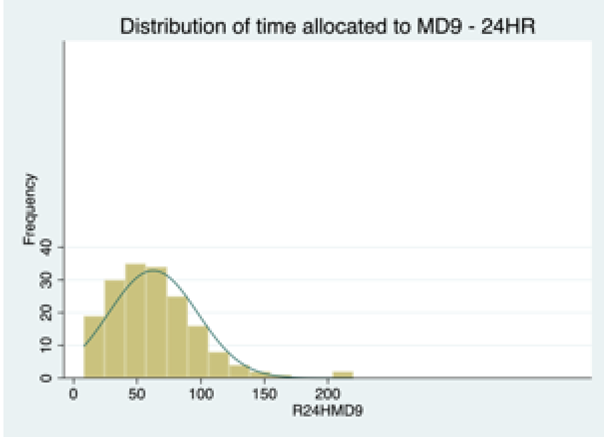
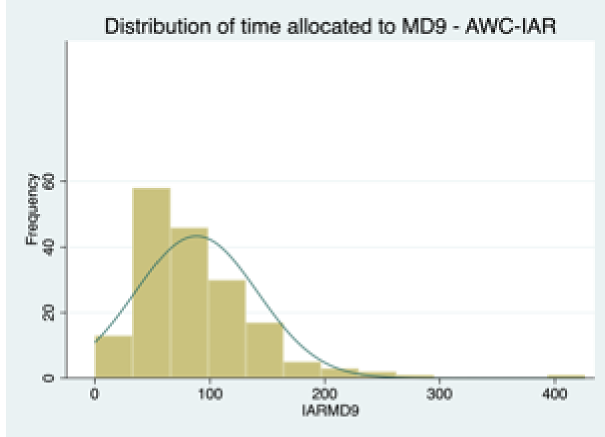
OBS, observation; 24HR, 24-hour recall; AWC-IAR, automated wearable camera-based image-assisted recall.

Additional Figure 6. Inter-method comparison of the distribution of time allocated to culture, leisure, mass media and sports practices.

OBS	24HR	AWC-IAR
		

OBS, observation; 24HR, 24-hour recall; AWC-IAR, automated wearable camera-based image-assisted recall.

Additional Figure 7. Inter-method comparison of the distribution of time allocated to self-care and maintenance.

<table>
<tr><td>OBS</td><td>24HR</td><td>AWC-IAR</td></tr>
<tr><td>
</td><td>
</td><td>
</td></tr>
<tr><td colspan="3">OBS, observation; 24HR, 24-hour recall; AWC-IAR, automated wearable camera-based image-assisted recall.</td></tr>
</table>

Chapter 7: Thesis discussion and conclusions

In this final chapter, the main findings from the research paper chapters of this thesis (Chapters 4-6) will be integrated and summarized as they relate to the primary objectives of the thesis. Additionally, the strengths and limitations of the study design, the relevance of the findings as they relate policy and program considerations, and areas for future relevant research will be discussed.

MAIN FINDINGS

A systematic literature review (Chapter 2) showed there was insufficient evidence to determine the feasibility, acceptability, or validity of the AWC-IAR or 24HR methods for assessing maternal or child dietary diversity or women's time use, in rural LICs, and that the evidence gaps were substantial.

To my knowledge, no AWC feasibility or acceptability studies have been completed and published in a LIC. Among AWC feasibility and/or acceptability studies conducted in other (non-LIC) contexts, none have targeted women, mothers of children under two years of age, or people living in rural areas.

There were no criterion validation studies of the AWC-IAR method for assessing dietary diversity in any context, and only one criterion validation study of the 24HR method for assessing dietary diversity among women of reproductive age, which, consistent with our study, found that 24HR methods (list-based and open recall) overestimated the DDS and MDD-W compared to WFRs. (107,131) There were no criterion validation studies of either the AWC-IAR or 24HR methods for assessing infant and young child DDS.

For assessing time-use, there were no criterion validation studies of the 24HR method in any context and no AWC-IAR criterion validation studies in LICs. There were also no criterion validation studies of either the AWC-IAR or 24HR methods for assessing time use among mothers of infants and young children.

The absence of studies demonstrating the validity of 24HR methods for assessing dietary diversity or time use is surprising given their widespread use, globally. These findings underscore the importance of the research undertaken in this thesis. The main findings of this thesis are summarized in Table 7.1.

Table 7.1. Thesis main findings

Objective		Main Findings
1	Assess the feasibility of an automated wearable camera-based image-assisted recall for use in collecting dietary and time-use data in rural Eastern Region Uganda.	• Most study participants reported their experiences with the AWC and IAR to be good or very good and would participate in a similar study again. • None of the eight study withdrawals could be definitively attributed to AWC acceptability, but three occurred because the participant's spouse refused to participate after the participant had consented. • AWC acceptability appeared to have varied between the two study districts. • Providing participants time to review their AWC-captured images in private before being viewed by the research team, a practice promoted to ensure study participants' ethical rights, was infeasible in this context, due to low-literacy and limited exposure to digital technologies. • Fifteen percent of data was lost due to AWC malfunction, and twelve percent of the AWC images were "uncodable" due to insufficient lighting. • The manual methods used to process and analyse the AWC data in this study, were labour-intensive, time-consuming, and prone to human error. • Some participants had difficulty interpreting first-person perspective photographs.
2	Evaluate the concurrent criterion validity, for assessing maternal and child dietary diversity, using a semi-quantitative automated wearable camera-based image-assisted recall method and a semi-quantitative multiple-pass 24-hour dietary recall method, compared to weighed food records in rural Eastern Region Uganda.	• Systematic bias was low for both 24h dietary recall and AWC-IAR, indicating that both methods provide an accurate estimate of the median DDS in this rural Ugandan context. • Random error was high, indicating that DDS estimated via 24HR and AWC-IAR were unreliable at the individual level. Estimates of DDS could differ by more than two food groups from observed values, and over a third of DDS were misclassified, • BAPs showed a consistent and uniform pattern across the range of mean DDS, indicating that 24HR and AWC-IAR methods performed equally well for participants at both the lower and higher ends of the DDS spectrum.
3	Evaluate the concurrent criterion validity, for assessing women's time use, of an automated wearable camera-based image-assisted recall method and a 24-hour activity recall method, compared to direct observation in rural Eastern Region Uganda.	• For both the 24h activity recall and AWC-IAR methods, systematic bias was low for women's time allocation to employment, own production, domestic chores, and self-care, indicating that both methods provide an accurate estimate of median time allocated to these activities in this rural Ugandan context. • Women's time allocations to caregiving and socializing were underestimated by both methods. • For most activities, random error was high (>2 hrs), most notably for caregiving and socializing (>7 hrs), indicating estimates of time allocated to these activities via 24HR and AWC-IAR were unreliable at the individual level. • The time allocation BAPs showed varying patterns across activity groups.

		• Only the IAR method generated cloud-shaped plots (domestic chores, caregiving, and socializing), indicating the method performed equally well for women spending little time doing these activities and women spending substantial time doing these activities. • Both the 24HR and the IAR methods had fan-shaped plots for employment, own production (IAR only), and leisure, indicating the amount of random error increased as the mean time allocated to the activity group increased. • The 24HR method had several downward-sloping BAPs (own production, domestic chores, caregiving, and self-care), and IAR had one downward-sloping plot for self-care, indicating that, for these activities, the methods underestimated time allocations for women on the lower end of the spectrum and overestimated time allocated to the activity for women at the upper end of the spectrum.

In this thesis, I investigated aspects of AWC feasibility and acceptability from the perspectives of participants and the research team. (See Chapter 4) Collection of food / beverage intake and women's time-use data, using an AWC, was acceptable and feasible. Most participates rated their experiences with the AWC-IAR as either good or very good and would be willing to participate in a similar study again in the future.

There were no reports of physical discomfort due to the AWC. There also were no instances of lost data due to participants forgetting to wear the AWC or the researcher failing to recharge it. Nor were any usability issues reported. We aimed to minimize interaction between the participant and the AWC to avoid the usability-related loss of data seen in previous studies. To reduce excessive movement of the device, as noted in other studies wherein participants wore the AWC on a lanyard around their neck, a member of the research team (Jan Priebe) custom-designed a clip that securely fixed the AWC to an article of clothing. To standardize the AWC placement at the neckline, the device was clipped to a study-provided t-shirt worn by the participant over their usual clothing. Instead of handling the AWC, participants were instructed to remove the t-shirt when they wanted privacy. AWCs were delivered and affixed to study participants at the start of the data collection day and collected from the participants at the end of the data collection day and charged overnight by a member of the research team (ALSB).

The difficulty some participants had when interpreting first-person perspective photographs was an unanticipated feasibility challenge. The AWC-IAR method is designed so that images prospectively capture by the AWC can be used later in an IAR to "trigger" participants memories. It is therefore notable that half of participants had difficulty interpreting at least one of the AWC orientation images. Among the five image types included in the orientation, the photo of the participant's own hand while performing a task was the most problematic.

We were also surprised to find that the enumerator-only image interpretation (i.e., the enumerator coding the images alone, based on what she could see in the images, without the assistance of the participant) performed poorly. Although several previous studies have used an enumerator image interpretation approach, (28,29,31,33,37,67,68) we would conclude from this study that engagement from the participant is required to accurately interpret the foods / beverages consumed and activities performed.

Although eight participants voluntarily withdrew from the study, none of these could be definitively attributed to AWC acceptability. Three of the eight withdrawals occurred because the participant's spouse refused to participate after the participant had consented, which underscores the importance of careful sensitization of not only potential AWC-based study participants but also their families and their communities. Another unanticipated finding was that a higher proportion of the participants withdrew from one study district compared to the other (i.e., 9% in Kamuli vs 1% in Bugiri), which may indicate that acceptability can vary across small scale geographies. Data collection in Kamuli occurred in the last two weeks of the intensive six-week study. During this period, the field coordinator (ALSB) became severely ill with typhoid, which stretched the capacities of the remaining team members. It is therefore also possible that the field team, fatigued after a month of 15-day observations and overstretched, may have lacked the bandwidth to adequately attend the needs of study participants during observation or while facilitating the recalls.

It is common practice in AWC-based research in HICs to allow participants time to review their images in private before being viewed by the research team, and to delete any images desired. This approach, however, is impractical in a low-literacy population with limited exposure to digital technologies. We attempted to overcome this challenge by asking the participant at the end of each data collection day, before images had been viewed by the participant or the researcher, whether they would like us to delete all images or, with the assistance of the researcher, to delete some selected images. Participants were also invited to request deletion of any or all images as they viewed them with the researcher during the IAR. Despite being repeatedly offered the opportunity to do so, none of the participants requested that their data be entirely deleted either at the end of the data collection day or after viewing their photos during the IAR. We cannot determine, with the available data, whether the lack of any requests to delete all photos occurred because the participants found the content of the AWC images acceptable or, instead, that they did not feel comfortable asking the enumerators for the photos to be deleted. Seven participants did request for a few images to be deleted, however, suggesting that the protocol was at least acceptable for some.

Among the most grievous feasibility challenges was the loss of data due to AWC malfunction. Twenty-seven participant days were lost due to AWC inoperability, and twelve percent of images were too dark to interpret. These issues have important cost implications for future AWC-based research study designs, either for recruiting more participants in anticipation of lost data and / or procuring higher quality devices. A larger sample size will not, however, address the problem of poor AWC image quality due to low-light conditions. In this study, for the purpose of method validation, we restricted the period of comparison to daylight hours. We also determined that neither foods groups consumed nor activities performed before 8am or 8pm were different than those during the period of comparison. However, for future stand-alone AWC studies, the results may be biased if the outcome of interest occurs more or less frequently in low-light conditions. In these cases, new solutions are needed for artificially lighting the activity space.

Processing and analysing the AWC data was labour-intensive, time-consuming, and prone to human error. The time and expertise required to process and analyse AWC-derived data is a major obstacle for adoption in large-scale studies. Integrated software for simultaneous viewing and coding of AWC-captured images that work on tablet computers could potentially decrease the burden for research teams. AI is another potential labour-saving option. AI has been used to determine portion size from food photos in lab conditions, however, to the best of my knowledge the approach has not yet been attempted on AWC-captured images from participants in free-living conditions. (146–148)

The AWC-IAR performed similarly to 24HR in assessing dietary diversity for mothers and children. (See Chapter 5) Both the 24h dietary recall and AWC-IAR methods provided an accurate estimate of the median DDS for women and children, as illustrated by the low systematic bias and consistent cloud-shaped pattern, in the BAPs. These results indicate that both the 24h dietary recall and the AWC-IAR methods are valid for assessing median maternal and child DDS at the population level. The extent of the random error for both methods, however, was high. Estimates of DDS could differ by more than two food groups from observed values and over a third of DDS were misclassified. The BAPs showed a consistent and uniform pattern across the range of mean DDS, which indicates that both methods perform similarly for women and children across the range from low to high DDS.

The results of the analyses of the AWC-IAR and 24h activity recall methods were different. (See Chapter 6) For both the 24h activity recall and AWC-IAR methods, systematic bias was low for women's time allocation to employment, own production, domestic chores, and self-care. For these activities, both the 24h activity recall and the AWC-IAR methods are valid for assessing women's median time allocation at the population level. Women's time allocation to caregiving and socializing, however, was underestimated by both methods (>1 hr). Underestimation of time allocated to these activities was apparently higher for the 24h activity recall than the AWC-IAR (i.e., 226 m vs. 62 m for time allocated to caregiving and 172 m vs. 109 m for time allocated to socialising for 24h activity recall and AWC-IAR, respectively). For most activities, the random error, for both methods, was high (95% confidence limits >2 hrs). For caregiving and socializing, the 95% confidence limits for random error exceeded 7 hours. These results indicate that for most activities, neither the 24h activity recall nor the AWC-IAR are valid for estimating time use at the individual level.

The shape of the time-use BAPs varied by activity. The 24HR method had several downward-sloping BAPs (own production, domestic chores, caregiving, and self-care), and IAR had one downward-sloping plot for self-care, indicating that, for these activities, the methods underestimated time allocations for women on the lower end of the spectrum and overestimated time allocated to the activity for women at the upper end of the spectrum. Both the 24HR and the IAR methods had fan-shaped plots for employment, own production (IAR only), and leisure, indicating the amount of random error increased as the mean time allocated to the activity group increased. Only the IAR method generated cloud-shaped plots (domestic chores, caregiving, and socializing), indicating the method performed equally well for women spending little time doing these activities and women spending substantial time doing these activities. That only the IAR method resulted in BAPs exhibiting the ideal shape (cloud-shaped) suggests that, for assessing time allocation to domestic chores, caregiving, and socializing, the AWC-IAR method may result in more predictable bias compared to 24h activity recall method.

METHODOLOGICAL CONSIDERATIONS

Strengths

This thesis comprises the first studies, to our knowledge, to have rigorously assessed the feasibility and acceptability of AWCs, and the concurrent validity for assessing maternal and child dietary diversity and women's time use, of an AWC-IAR method and 24HR methods in a rural LIC context. "Gold standard" reference methods were used for dietary diversity (WFRs) and time use (OBS). The Bland-Alman LOA method was used for data analysis, which allowed us to assess both the systematic bias and random error, and how it changes across the spectrum from low to high DDS and/or time allocation.

This study was conducted in communities with free-living participants which reflects conditions in which research is typically carried out in rural LICs. Compared to previous AWC feasibility studies, the majority of which (88%) were conducted with populations less than a hundred, this study was conducted with a relatively large number of participants.

It also reports on several research design innovations that were used to adapt AWC-based research protocols commonly used in HICs to a LIC context. First, participants in this study were not responsible for operating the AWC devices, which, although increasing the burden on the research team, likely reduced lost data. Second, an orientation to the AWC-captured images was added at the start of the IAR to help participants with little previous experience viewing first-person photography to use the images to trigger their memories. Third, the IAR protocol designed for this study did not require the participant to operate the computer. Instead, a touchscreen tablet was used so that the participant and enumerator could review the images together, while the enumerator simultaneously recorded the foods / beverages consumed and activities performed. Furthermore, whereas dietary diversity and time-use data were collected using separate 24HR instruments, the AWC-IAR protocol we designed integrated data collection in a single instrument. This reduced repetitiveness compared to successive 24HRs (i.e., one for dietary diversity and one for time use) and, by simultaneously capturing of a behaviour of interest (i.e., dietary intake) and the social contexts in which it is performed (e.g., position in a series of activities, location, other individuals involved) may have helped to improve the quality of the information obtained. Finally, having the enumerator first code participants' images captured by the AWC before reviewing the images with the participant enabled us to assess whether the mothers' participation in image interpretation and analysis is necessary. It also reduced the time required to administer the IAR because the enumerator could direct the interview to image segments requiring clarification rather than scrolling through all images one-by-one.

Limitations

This study was conducted in a single population and so the results cannot be generalized to other populations, even to other rural populations in Uganda. In rural contexts, seasonality has a heavy influence on diets, due to the fruits and vegetables that are available, and time use, due to the agricultural work cycle. Therefore, the results of this study, undertaken during the dry season, cannot be generalised to other seasons. The acceptability of the IAR, which took approximately 1-2-hours to administer, may be lower in periods when workloads are heavy, and women are experiencing time constraints e.g., during the planting or harvest seasons.

The feasibility of an AWC may be lower for participants who work away from home, for a myriad of reasons including that participants may be more likely to be forbidden by formal employers from wearing the device; the implications of forgetting to wear the AWC may be more severe if participants cannot easily and quickly access the device; active or passive interactions with a higher number of members of the public may expose the participant to more negative reactions; and, if the work performed by commuting participants is indoors, for example in a large market, poor lighting may result in a higher proportion of images that are too dark to code.

Different foods or diets may be more or less difficult to identify with the AWC captured images. The foods consumed by participants in this study were almost exclusively prepared at home and simple, and diets were fairly consistent across participants. More foods may be missed by the AWC in contexts where snacking is common. It may also be difficult to objectively determine the food groups consumed in contexts where foods prepared outside the home comprise a large proportion of the diet. The foods consumed may be more reliably captured by an AWC worn at the neckline in context where eating primarily occurs at a table, rather than on the ground.

Different activities may also be more or less difficult to identify with the AWC captured images. Participants in this study spent most of their time at home. The validation results may be different in a population in which a larger proportion of women spent more time in employed work, where the activities performed may be less varied, making coding simpler, especially if they are engaged in more formal work, such as administrative support in an office. However, activities commonly performed while commuting, such as reading, chatting, or listening to music on a mobile phone, may be missed.

Because the feasibility study was nested within a rigorous validation study, which necessitated the concurrent administration of multiple methods, the methods used to assess their perceptions of the AWC and IAR were somewhat short-changed. More in-depth qualitative methods (such as focus group discussions with participants, and their families and communities) could have helped to better understand the surprisingly positive favourability ratings obtained.

The extensive interaction between the study team and participants, due to the concurrent collection of data via multiple methods (e.g., 15-hour OBS, 24HRs and IVR), may have biased the feasibility results in several ways. For example, the number instances of lost data due to inoperability was higher on days when the observer was not present at the home (n=24 and n=32 when the observer was and was not present, respectively). Lost data was possibly lower in this study than it would be in a stand-alone AWC-based study because an observer in the home would have more readily identified a malfunctioning AWC and alerted the research team for quick replacement. Acceptability-related study withdrawals may have been lower in this study because the observers helped to monitor negative rumours so the research team could address them early.

Conversely, acceptability may have been adversely affected because of high participant burden due to multiple simultaneous methods. For example, on the day after observation day, four separate recall methods were administered to the participant in succession (24h dietary recall for children, 24h dietary recall for women, 24h activity recall for women, and the IAR). The AWC-IAR method itself concluded with a dietary diversity checklist for women, dietary diversity checklist for women, similar to those used in the respective 24HR methods, and women's activities checklist. The intensive succession of seemingly repetitive questions likely posed an especially heavy burden on both participants and interviewers that is not typical of most studies, which may have contributed to poorer performance on the AWC-IAR than would otherwise have been the case. Resource constraints prohibited administering the 24HR methods on a different day than the AWC-IAR.

Whereas the OBS and 24h activity recall protocols were both closed-ended and shared a similar matrix structure of pre-defined activity categories, the AWC-IAR structure was open-ended. Keeping a similar structure to the other methods would have posed less burden on the field enumerators and data entry RAs. Had resources been available, a CAPI (Computer-Assisted Personal Interviews)-based data collection tool (compared to paper-based data collection) that, for example, prompted enumerators when a series of recorded activities was incomplete, may have further reduced the burden on field enumerators, the amount of data that needed to be entered by the RAs ex post facto and error related to incomplete data capture.

Coding the time-use data was substantially more challenging than the dietary intake data. Although the field enumerators and data entry RAs received the same training in the coding of activities, there were variations in how the same or similar activities were coded. Inconsistencies in the coding of activities within activity groups posed no problem, since the time allocation was analysed at the (aggregated) activity group level. For example, playing with a young child and feeding a young child are somewhat ambiguous activities, but both fell under the caregiving activity group. Underestimations in one activity would have offset overestimates in another activity within the same activity group. Due to the balancing out that occurs when time allocated to activities are aggregated, the results may vary depending on the level of aggregation and composition of the activity groups. Time-use results can only be compared (e.g., within the same population over time or between studies) when a standard lexicon (e.g., ICATUS or HETUS) is used.

A different enumerator conducted the observation than administered the recall methods (i.e., 24HRs and IAR). Between field enumerator differences in how the same activity was coded for activities in the same activity group would also have cancelled out when aggregated for comparison. Between enumerator coding difference for activities in different activity groups, however, would have been more problematic. If the recall enumerator was always paired with the same observation enumerator, between-enumerator coding difference for activities in different activity groups would have appeared as systematic errors. If the observation / recall enumerator pairings were randomized, then between-enumerator coding difference for activities in different activity groups would have appeared as random errors. Future validation studies should randomize the recall enumerator to reduce study-induced systematic errors.

OBS and 24HR methods were coded in the field, using a predefined list of activities. The IAR, however, was captured in the field using an open format, which was then later coded (by a different team of RAs) using the lexicon used for OBS and 24HR methods. A few activities were coded into different activity groups by field enumerators and data entry RAs. For example, peeling sweet potatoes and shelling and pounding groundnuts were variably coded as food preparation (domestic chores) or post-harvest processing (own production). To address this issue, post-harvest processing activities were remapped to the domestic chores activity group. While this issue was caught, other coding discrepancies may have contributed to the random error seen for the AWC-IAR in this study. Formative research prior to undertaking a time-use study would also help to identify activities commonly performed by the study population and ensure similar activities are well defined for the research team.

There are a few aspects of acceptability that this study did not address. The possible effect of the AWC on recruitment was not investigated. Also, our study did not assess acceptability from the perspective of participants' family members or community, who might be incidentally caught on camera. In-depth qualitative research would have been helpful to better understand these and other considerations regarding AWC acceptability for participants and their communities. Further research is also needed to estimate the cost-effectiveness of the AWC-IAR method versus traditional recall methods, accounting for all equipment costs (including but not limited to the AWCs) and researcher time required.

Whereas study enumerators had prior experience performing OBS and facilitating 24HRs (other data collection methods assessed in this study), the IAR was a new and complex method that required strong verbal probing skills and attention to detail. Although enumerators were provided guidance and training in how to facilitate an IAR, there was no predetermined script provided. Enumerators had to figure out how to elicit the information from each participate based on the images before them on the tablet. In training, we found that some enumerators picked up the IAR technique easily whereas other enumerators really struggled. Future AWC-IAR studies should ensure adequate time for enumerator training, evaluation of enumerator IAR facilitation skills prior to field work, and supportive supervision during data collection.

Research and Program Impact Evaluation Implications

I hypothesized that an AWC-IAR method for assessing women's and children’s dietary diversity and women's time was feasible in a rural LIC context and could provide accurate and reliable data compared to OBS. The results of this thesis indicate that AWCs are indeed feasible and acceptable for research data collection in rural LIC contexts, although challenges remain.

The AWC-IAR method provided valid results for maternal and child DDS and MDD at the population level and performed well for women and children at lower and higher ends of the DDS spectrum. Prior to this study, it was known that, in quantitative dietary assessment, measurement error (the difference between reported energy and nutrient intakes and true intakes) commonly occurs. This study showed that a high degree of measurement error also occurs when diet quality is assessed by the number of food groups consumed. That the AWC-IAR method is inappropriate for assessing DDS at the individual level (as indicated by high random error) is true also for 24h dietary recall methods and, for this reason (among others), neither 24HR- or AWC-IAR-derived DDS should be used to assess associations between dietary diversity and other outcomes of interest (e.g., stunting, overweight / obesity, or other diet-related non-communicable diseases). Given their similar performances when compared to WFR, with regard to validity, the AWC-IAR may be a viable option to the 24HR method for researchers seeking to assess maternal and/ or child dietary diversity in rural, low-income contexts.

For assessing women's time use, the results of this thesis show that neither the AWC-IAR nor the 24h activity recall methods are equally valid at the population or individual levels across all activities. In particular, both the AWC-IAR and 24h activity recall methods perform most poorly for activities commonly performed by women in rural LIC contexts (caregiving and socialising). The implications of these findings are perhaps the most urgent for the 24h activity recall method. The WEAI and LSMS, for example, use similar but different lexicons than the ICATUS used in this study. Given the poor results obtained in this study for assessing some activities via the 24HR method, and the widespread use of the WEAI and LSMS time-use modules across low- and middle-income countries, it seems prudent to conduct similar validation studies using these protocols.

Prior to this study, it was known that traditional time use assessment methods tended to underestimate women's time use and did not adequately account for multi-tasking. This study showed that 24h activity recall may underestimate women's time allocated to caregiving up to 60%.

For dietary diversity, the AWC-IAR method and 24HR methods performed similarly. For time allocations to domestic chores, caregiving and socializing especially, the results of this study suggest that time-use estimates obtained via the AWC-IAR method are more accurate and reliable than those obtained via the 24HR method. The BAP shapes indicate that, via the 24h activity recall, as the mean time allocated to these activities increased, the magnitude and/or direction of the error appeared to change. Overall, however, this thesis reveals the need for new thinking regarding time-use assessment methods because, for most activities, neither method performed well.

The benefits of the AWC-IAR method must be carefully weighed against the additional burdens imposed on the research team and potential for data losses. Although issues related to AWC malfunction, poor image quality, and poor device usability seen in this study may be alleviated by an investment in higher quality cameras, these challenges are commonly reported in other AWC studies using top-of-the-line devices explicitly designed for the purpose of behavioural research data collection. Moreover, the cost of the devices is only one of the costs incurred when conducting an AWC-based study. The true cost of the AWC-based method must also account for, among other things, the cost of tablets to view the images, peripherals to transfer and store the data, and either bespoke software and/or researcher time to analyse the data. When comparing two equally valid methods, as is the case for the 24HR and AWC-IAR methods in assessing dietary diversity, comparative cost-effectiveness studies that account for all aspects of the methods are needed in order to assess the trade-offs. In the immediate future, AWS-based studies should over-recruit in anticipation of lost data due to malfunctioning equipment. Consideration must also be given to the reference period to ensure adequate lighting, and/or innovative solutions to artificially light the activity space of participants.

A final important implication of this thesis regards research ethics. Low literacy and limited exposure to computers among residents of most rural LIC areas necessitate adjustments to AWC ethics protocols commonly used in HIC contexts. Kelly et al.'s (2013) AWC ethical framework is based on the Declaration of Helsinki (1964) and Nuremberg Code (1947) principles of respect for autonomy, beneficence, non-maleficence, and justice. (106) Among other ethical guidelines, it calls for participants of AWC-based studies to "have the opportunity to view (and delete if necessary) their images in privacy". In rural LIC contexts, however, participants may not have the skills to review their AWC-captured images independently. As social norms and individual capacities may differ in and among LICs, other aspects of the framework may also warrant review for appropriateness and applicability.

FUTURE RESEARCH

This thesis marks an important step in understanding the feasibility and validity of an AWC-IAR method in a rural LIC context. Together, the three sets of analyses are the first evidence for: the feasibility and acceptability of AWC-based research methods among women, mothers of infants, and/or persons living in a rural LIC area; the feasibility of AWC-based research in rural LIC contexts for the research team; and the validity of an AWC-IAR method in a rural LIC context for assessing maternal and child dietary diversity and women's time use. The thesis, as far as I'm aware, also comprises the first criterion validation of a 24HR method for assessing time use in any country, and the second for assessing dietary diversity in a LIC. The thesis therefore provides a strong foundation for future research regarding AWC-based data collection methods generally and AWC-IAR methods specifically in LICs, and AWC-IAR and 24HR methods for assessing maternal and child dietary diversity and women's time use in rural LIC contexts. Below I summarize my recommendations for the use of AWC-based research in LICs based on the findings of this thesis.

Although the AWCs were found to be feasible for participants in this study, further work is needed to improve the feasibility of AWC-based methods for researchers. First, a review of the applicability of the Kelly et al. (2013) AWC ethics framework by LIC stakeholders is urgently needed. (106) Researchers from LICs, representatives from government and/or IRBs, potential study participants, local leaders, and other members of civil society should be convened for a multi-stakeholder workshop to identify the gaps and concerns, and to adapt (or create) a framework that both adheres to the key ethics principles of respect for autonomy, beneficence, non-maleficence, and justice and is applicable to typical LIC norms and constraints. The inability of a large portion of the population in LICs to reliably operate a computer for the purpose of independently reviewing AWC-captured images is one obvious example. The Kelly framework also seems to assume a robust policy environment, where laws regarding e.g., the use of recording devices by the public and the right to personal privacy are in place and observed, which is not always the case for LICs.

In order for AWC-based research to be conducted at scale, robust field tools that reduce data losses and the time required to process and code AWC images are needed. This includes new solutions for illuminating the activity space when electricity is unavailable or cost-prohibitive (for participants), a CAPI for collecting IAR data in the field, and improved protocols for efficiently reviewing and coding the images with participants. Notably, AWCs are now no longer being manufactured, having been replaced by wearable video cameras. These devices are generally more expensive than AWCs, and also generate far more images for the same amount of data collection time, imposing additional storage challenges and burden on the research team for data analysis. The time is ripe for new thinking regarding the use of these tools in research. For example, it may be easier and faster for the participant to review a lower number of representative photos pre-selected by the enumerator rather than scrolling through all 1,500 images, or to review the images as a video rather than individual pictures.

Also, future studies may explore opportunities for reducing the labour involved in analysing AWC-captured data. In HIC contexts, researchers are using AI to automatically identify and/or quantify foods (147,149–152) and determine activity-related information (153–155) using images captured via an AWC. These studies are primarily undertaken in controlled environments, but perhaps there is an opportunity to use AI to assist image coding in LIC contexts too. Although AI may not be able to replace a human researcher, it seems plausible that it could be used to substantially reduce the number of photos a researcher must inspect. For example, AI may not be able to identify a single mango in a series of 1,500 images, but it may be able to identify series of photos where foods are being consumed.

Additional AWC feasibility and acceptability studies are needed in other rural, peri-urban, and urban settings of LICs, where time use and dietary patterns, levels of education and/or exposure to technology, for example, might be quite different. Future studies should carefully track the participation rate (that is, the proportion of individuals who choose to participate in the study out of the total number invited) and, where possible, the reasons for their decision. More work is also needed to assess acceptability from the perspective of the public. Future studies should investigate the analytical benefits of AWCs unique capacity to simultaneously capture data on a variety of individual-, household- and community-level factors that influence maternal and child health and nutrition, in comparison with standardized stand-alone tools. Studies are also needed to assess the validity of the AWC-IAR method for measuring behaviour change over time, for example before and after an intervention.

This thesis reveals that both the AWC-IAR and 24HR methods are poor for estimating women's time allocated to caregiving. Even when viewing photos of themselves supervising a child, study participants often neglected to report caregiving as an activity. The WEAI time-use module, a 24HR method, has recently been modified to explicitly query women's caregiving responsibilities at every timeslot. To the best of my knowledge, however, no before- or after-validation studies have been conducted to determine if the change has resulted in more accurate and reliable reporting. Future AWC-based studies might explore techniques for helping time use study participants to take notice of caregiving and other "taken for granted" activities. (140)

Finally, although the BAPs make visible changes in the extent and/or direction of measurement error across the spectrum from low to high of the measured outcome, we do not know the implications of the three BAP shapes seen in this study (cloud, fan, and downward sloping) on attenuation or how attenuation may vary depending on the outcome measured. Further research is needed to model the error for different shaped BAP patterns and outcomes of interest.

CONCLUSIONS

This thesis comprises the first-ever AWC feasibility and validity research conducted in a LIC, and important findings regarding traditional (24HR) methods. The research findings indicate that an AWC-IAR is a feasible method for collecting maternal and child dietary diversity and women's time-use data in a rural LIC context, although challenges remain. The AWC-IAR and 24HR methods provided an accurate estimate of median maternal and child dietary diversity and minutes women spent in employment, own production, and domestic chores. The median time women spent caregiving and socializing, however, was underestimated by both methods, and, for these activities, the extent of the underestimation via the 24h activity recall was apparently higher than via the AWC-IAR. High LOA for dietary diversity and time allocation indicated that associations between either dietary diversity or time use and outcomes would be attenuated. For dietary diversity, the BAPs for both methods showed a consistent and uniform pattern across the range of mean DDS. The consistent BAPs of the AWC-IAR compared with the fan-shaped or downward trend shaped BAPs of the 24 hour activity recall method indicate that the AWC-IAR method for assessing time allocation to domestic chores, caregiving, and socializing may result in more predictable bias compared to 24h activity recall method.

Currently, however, AWC-based research methods pose a heavy burden on researchers, which must be considered among the trade-offs when designing future studies. For DDS, the validity of the AWC-IAR method was similar to the 24HR method, whereas the time required to analyse the data was substantially higher. With the current set of tools available, the added value of the AWC-IAR may not outweigh the resource costs for the research purpose. Conversely, for time-use, the validity of the AWC-IAR method was substantially better than the 24HR for several activity groups in this study context. The case for using the AWC-IAR method for future time use studies is much stronger.

To be a more viable method for collecting data at scale in LICs, improvements are needed in study protocols and data collection tools. In particular, the AWC ethical framework commonly used in HICs needs to be reviewed for applicability to LIC contexts, new solutions are needed to overcome challenges posed by the limited access to electricity, and protocols and tools for coding and analysing AWC-captured images that work in LIC contexts and reduce participant and researcher burdens are needed. Nonetheless, AWCs offer a unique opportunity for simultaneously capturing data at individual, family, and community levels that might otherwise go unreported by study participants because it never made an impression.

References

1. Black RE, Victora CG, Walker SP, Bhutta ZA, Christian P, De Onis M, et al. Maternal and child undernutrition and overweight in low-income and middle-income countries. Lancet [Internet]. 2013;382(9890):427–51. Available from: http://linkinghub.elsevier.com/retrieve/pii/S014067361360937X

2. Stanaway JD, Afshin A, Gakidou E, Lim SS, Abate D, Abate KH, et al. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Stu. Lancet [Internet]. 2018 Nov;392(10159):1923–94. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673618322256

3. Afshin A, Sur PJ, Fay KA, Cornaby L, Ferrara G, Salama JS, et al. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2019 May;393(10184):1958–72.

4. Horton S, Steckel RH. Malnutrition: Global economic losses attributable to malnutrition 1900-2000 and projections to 2050. Copenhagen: Copenhagen Consensus on Human Challenges; 2011. (Copenhagen Consensus on Human Challenges Assessment Paper).

5. FAO. The State of Food Insecurity in the World 2013: The multiple dimensions of food security. Rome: Food and Agriculture Organization of the United Nations; 2013. 114 p.

6. United Nations Children’s Fund (UNICEF), World Health Organization, International Bank for Reconstruction and Development/The World Bank. Levels and trends in child malnutrition: Key findings of the 2021 edition of the joint child malnutrition estimates. Geneva: World Health Organization; 2021.

7. UNICEF. Strategy for improved nutrition of infant and women in developing countries. New York: UNICEF; 1990.

8. Black RE, Allen LH, Bhutta ZA, Caulfield LE, de Onis M, Ezzati M, et al.

Maternal and child undernutrition: global and regional exposures and health consequences. Lancet [Internet]. 2008;371(9608):243–60. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0140673607616900

9. World Health Organization (WHO), United Nations Children's Fund (UNICEF). Indicators for assessing infant and young child feeding practices: definitions and measurement methods. Geneva: World Health Organization and the United Nations Children's Fund (UNICEF); 2021.

10. Engle PL, Menon P, Haddad L. Care and Nutrition: Concepts and Measurement. Washington, D.C.; 1997.

11. International Labour Office. Women at Work: Trends 2016. Geneva; 2016.

12. Kes A, Swaminathan H. Gender and Time Poverty in Sub-Saharan Africa. In: Blackden CM, Wodon Q, editors. Gender, Time Use, and Poverty in Sub-Saharan Africa. Washington, D.C.: The International Bank for Reconstruction and Development/The World Bank; 2006. p. 13–38.

13. Ricci J, Jerome N, Megally N, Galal O, Harrison G, Kirksey A. Assessing the Validity of Informant Recall: Results of a Time Use Pilot Study in Peri-Urban Egypt. Hum Organ [Internet]. 1995;54(3):304–8. Available from: http://sfaajournals.net/doi/10.17730/humo.54.3.a536540628381010

14. Robinson JP. The time-diary method: Structure and uses. In: Pentland WE, Harvey AS, Lawton MP, McColl MA, editors. Time Use in the Social Sciences. New York: Kluwer Academic Publishers; 1999. p. 47–89.

15. Gross DR. Time allocation: A tool for the study of cultural behavior. Annu Rev Anthropol. 1984;13:519–58.

16. Bério A-J. The analysis of time allocation and activity patterns in nutrition and rural development planning. Food Nutr Bull [Internet]. 1984 Mar 2;6(1):1–16. Available from: http://journals.sagepub.com/doi/10.1177/156482658400600113

17. Marini MM, Shelton BA. Measuring Household Work: Recent Experience in the United States. Soc Sci Res [Internet]. 1993 Dec;22(4):361–82. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0049089X83710185

18. Kalton G. Sample design issues in time diary studies. In: Juster FT, Stafford

FP, editors. Time, Goods, and Well-Being. Ann Arbor, Michigan: Institute for Social Reearch, The University of Michigan; 1985. p. 93–112.

19. Robinson JP. The Validity and Reliability of Diaries versus Alternative Time Use Measures. In: Juster FT, Stafford FP, editors. Time, Goods, and Well-Being. Ann Arbor, Michigan: Institute for Social Reearch, The University of Michigan; 1985. p. 33–62.

20. Wagenbuur HTM. An analysis of the time-budget and of the production and productivity of Lime Farmers In Southern Ghana. The Hague: Institute of Social Studies; 1972. (I.S.S. Ocassional Papers).

21. Gardner F. Methodological Issues in the Direct Observation of Parent–Child Interaction: Do Observational Findings Reflect the Natural Behavior of Participants? Clin Child Fam Psychol Rev. 2000;3:185–198.

22. Repp AC, Nieminen GS, Olinger E, Brusca R. Direct Observation: Factors Affecting the Accuracy of Observers. Except Child [Internet]. 1988 Sep 24;55(1):29–36. Available from: http://journals.sagepub.com/doi/10.1177/001440298805500103

23. Richer J. Direct Observation: Impediments and Approaches. Hum Ethol Bull [Internet]. 2017 Dec 28;32(4):6–14. Available from: http://ishe.org/human-ethology-bulletin/2017-2/heb-324-special-issue-why-direct-observation/direct-observation-impediments-and-approaches/

24. Bernard HR, Killworth P, Kronenfeld D, Sailer L. The Problem of Informant Accuracy: The Validity of Retrospective Data. Annu Rev Anthropol [Internet]. 1984;13(1):495–517. Available from: http://www.annualreviews.org/doi/10.1146/annurev.an.13.100184.002431

25. Beaton GH, Milner J, McGuire V, Feather TE, Little JA. Source of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. Carbohydrate sources, vitamins, and minerals. Am J Clin Nutr [Internet]. 1983 Jun 1;37(6):986–95. Available from: https://academic.oup.com/ajcn/article/37/6/986/4690872

26. Poslusna K, Ruprich J, de Vries JHM, Jakubikova M, van't Veer P, Vries JHM De, et al. Misreporting of energy and micronutrient intake estimated by

food records and 24 hour recalls , control and adjustment methods in practice. Br J Nutr [Internet]. 2018;101(2009):S73. Available from: http://www.journals.cambridge.org/abstract_S0007114509990602

27. Gibson RS, Charrondière UR, Bell W. Measurement Errors in Dietary Assessment Using Self-Reported 24-Hour Recalls in Low-Income Countries and Strategies for Their Prevention. Adv Nutr An Int Rev J [Internet]. 2017 Nov 15;8(6):980–91. Available from: http://advances.nutrition.org/lookup/doi/10.3945/an.117.016980

28. Arab L, Estrin D, Kim DH, Burke J, Goldman J. Feasibility testing of an automated image-capture method to aid dietary recall. Eur J Clin Nutr [Internet]. 2011;65(10):1156–62. Available from: http://www.nature.com/articles/ejcn201175

29. Arab L, Winter A. Automated Camera-Phone Experience with the Frequency of Imaging Necessary to Capture Diet. J Am Diet Assoc [Internet]. 2010 Aug;110(8):1238–41. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0002822310005304

30. O'Loughlin G, Cullen SJ, Mcgoldrick A, Connor SO, Blain R, Malley SO, et al. Using a Wearable Camera to Increase the Accuracy of Dietary Analysis. AMEPRE [Internet]. 2013 Mar;44(3):297–301. Available from: https://linkinghub.elsevier.com/retrieve/pii/S074937971200863X

31. Beltran A, Dadabhoy H, Ryan C, Dholakia R, Jia W, Baranowski J, et al. Dietary Assessment with a Wearable Camera among Children: Feasibility and Intercoder Reliability. J Acad Nutr Diet [Internet]. 2018 Nov;118(11):2144–53. Available from: https://doi.org/10.1016/j.jand.2018.05.013

32. Kamar M, Evans C, Hugh-Jones S. Factors Influencing British Adolescents' Intake of Whole Grains: A Pilot Feasibility Study Using SenseCam Assisted Interviews. Nutrients [Internet]. 2019 Nov 1;11(11):2620. Available from: https://www.mdpi.com/2072-6643/11/11/2620

33. Cowburn G, Matthews A, Doherty A, Hamilton A, Kelly P, Williams J, et al. Exploring the opportunities for food and drink purchasing and consumption by teenagers during their journeys between home and school: a feasibility study using a novel method. Public Health Nutr [Internet]. 2015 Jan

15;19(1):93–103. Available from: https://www.cambridge.org/core/product/identifier/S1368980015000889/type/journal_article

34. Gemming L, Doherty A, Kelly P, Utter J, Ni Mhurchu C. Feasibility of a SenseCam-assisted 24-h recall to reduce under-reporting of energy intake. Eur J Clin Nutr [Internet]. 2013;67(10):1095–9. Available from: http://www.nature.com/doifinder/10.1038/ejcn.2013.156

35. Beltran A, Dadabhoy H, Chen TA, Lin C, Jia W, Baranowski J, et al. Adapting the eButton to the Abilities of Children for Diet Assessment. In: Proceedings of Measuring Behavior 2016: 10th International Conference on Methods and Techniques in Behavioral Research International Conference on Methods and Techniques in Behavioral Research. Dublin, Ireland: NIH Public Access; 2016.

36. Gemming L, Rush E, Maddison R, Doherty A, Gant N, Utter J, et al. Wearable cameras can reduce dietary under-reporting: doubly labelled water validation of a camera-assisted 24 h recall. Br J Nutr [Internet]. 2015 Jan 28;113(2):284–91. Available from: https://www.cambridge.org/core/product/identifier/S0007114514003602/type/journal_article

37. Smith M, Stanley J, Signal L, Barr M, Chambers T, Balina A, et al. Children's healthy and unhealthy beverage availability, purchase and consumption: A wearable camera study. Appetite [Internet]. 2019;133(July 2018):240–51. Available from: https://doi.org/10.1016/j.appet.2018.11.004

38. Pettitt C, Liu J, Kwasnicki RM, Yang G-Z, Preston T, Frost G. A pilot study to determine whether using a lightweight, wearable micro-camera improves dietary assessment accuracy and offers information on macronutrients and eating rate. Br J Nutr [Internet]. 2016;115(01):160–7. Available from: http://www.journals.cambridge.org/abstract_S0007114515004262

39. Raber M, Patterson M, Jia W, Sun M, Baranowski T. Utility of eButton images for identifying food preparation behaviors and meal-related tasks in adolescents. Nutr J [Internet]. 2018 Dec 24;17(1):32. Available from: https://nutritionj.biomedcentral.com/articles/10.1186/s12937-018-0341-2

40. Raber M, Baranowski T, Crawford K, Sharma S V., Schick V, Markham C, et al. The Healthy Cooking Index: Nutrition Optimizing Home Food Preparation Practices across Multiple Data Collection Methods. J Acad Nutr Diet [Internet]. 2020 Jul;120(7):1119–32. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2212267220300629

41. Barr M, Signal L, Jenkin G, Smith M. Capturing exposures: Using automated cameras to document environmental determinants of obesity. Health Promot Int [Internet]. 2014;30(1):56–63. Available from: https://academic.oup.com/heapro/article-lookup/doi/10.1093/heapro/dau089

42. Gemming L, Doherty A, Utter J, Shields E, Ni Mhurchu C. The use of a wearable camera to capture and categorise the environmental and social context of self-identified eating episodes. Appetite [Internet]. 2015;92:118–25. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0195666315002469

43. McKerchar C, Smith M, Gage R, Williman J, Abel G, Lacey C, et al. Kids in a Candy Store: An Objective Analysis of Children's Interactions with Food in Convenience Stores. Nutrients [Internet]. 2020 Jul 18;12(7):2143. Available from: https://www.mdpi.com/2072-6643/12/7/2143

44. Mckerchar C, Smith M, Stanley J, Barr M, Chambers T, Abel G, et al. Food store environment examination – FoodSee: A new method to study the food store environment using wearable cameras. Glob Health Promot. 2019;0(0):1–9.

45. Schrempft S, van Jaarsveld CH, Fisher A. Exploring the Potential of a Wearable Camera to Examine the Early Obesogenic Home Environment: Comparison of SenseCam Images to the Home Environment Interview. J Med Internet Res [Internet]. 2017 Oct 12;19(10):e332. Available from: http://www.jmir.org/2017/10/e332/

46. Signal LN, Smith MB, Barr M, Stanley J, Chambers TJ, Zhou J, et al. Kids'Cam: An Objective Methodology to Study the World in Which Children Live. Am J Prev Med [Internet]. 2017 Sep;53(3):e89–95. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0749379717301630

47. Kelly P, Thomas E, Doherty A, Harms T, Burke Ó, Gershuny J, et al.

Developing a Method to Test the Validity of 24 Hour Time Use Diaries Using Wearable Cameras: A Feasibility Pilot. Harezlak J, editor. PLoS One. 2015 Dec 3;10(12):e0142198.

48. Gershuny J, Harms T, Doherty A, Thomas E, Milton K, Kelly P, et al. Testing Self-Report Time-Use Diaries against Objective Instruments in Real Time. Sociol Methodol [Internet]. 2020 Aug 30;50(1):318–49. Available from: http://journals.sagepub.com/doi/10.1177/0081175019884591

49. Harms T, Gershuny J, Doherty A, Thomas E, Milton K, Foster C. A validation study of the Eurostat harmonised European time use study (HETUS) diary using wearable technology. BMC Public Health [Internet]. 2019 Jun 3;19(S2):455. Available from: https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-019-6761-x

50. Davies A, Allman-Farinelli M, Owen K, Signal L, Hosking C, Wang L, et al. Feasibility Study Comparing Physical Activity Classifications from Accelerometers with Wearable Camera Data. Int J Environ Res Public Health [Internet]. 2020 Dec 13;17(24):9323. Available from: https://www.mdpi.com/1660-4601/17/24/9323

51. Doherty AR, Kelly P, Kerr J, Marshall S, Oliver M, Badland H, et al. Using wearable cameras to categorise type and context of accelerometer-identified episodes of physical activity. Int J Behav Nutr Phys Act [Internet]. 2013;10(1):22. Available from: http://ijbnpa.biomedcentral.com/articles/10.1186/1479-5868-10-22

52. Everson B, Mackintosh KA, McNarry MA, Todd C, Stratton G. Can Wearable Cameras be Used to Validate School-Aged Children’s Lifestyle Behaviours? Children [Internet]. 2019 Feb 1;6(2):20. Available from: http://www.mdpi.com/2227-9067/6/2/20

53. Gurrin C, Qiu Z, Hughes M, Caprani N, Doherty AR, Hodges SE, et al. The smartphone as a platform for wearable cameras in health research. Am J Prev Med. 2013;44(3):308–13.

54. Harvey JA, Skelton DA, Chastin SFM. Acceptability of novel lifelogging technology to determine context of sedentary behaviour in older adults. AIMS

Public Heal [Internet]. 2016;3(1):158–71. Available from: http://www.aimspress.com/article/10.3934/publichealth.2016.1.158

55. Kelly P, Doherty A, Berry E, Hodges S, Batterham AM, Foster C. Can we use digital life-log images to investigate active and sedentary travel behaviour? Results from a pilot study. Int J Behav Nutr Phys Act [Internet]. 2011;8(1):44. Available from: http://ijbnpa.biomedcentral.com/articles/10.1186/1479-5868-8-44

56. Kerr J, Marshall SJ, Godbole S, Chen J, Legge A, Doherty AR, et al. Using the SenseCam to Improve Classifications of Sedentary Behavior in Free-Living Settings. Am J Prev Med [Internet]. 2013 Mar;44(3):290–6. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0749379712008495

57. Smith C, Galland BC, de Bruin WE, Taylor RW. Feasibility of Automated Cameras to Measure Screen Use in Adolescents. Am J Prev Med [Internet]. 2019 Sep;57(3):417–24. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0749379719302065

58. Chambers T, Pearson AL, Kawachi I, Rzotkiewicz Z, Stanley J, Smith M, et al. Kids in space: Measuring children's residential neighborhoods and other destinations using activity space GPS and wearable camera data. Soc Sci Med [Internet]. 2017;193:41–50. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0277953617305907

59. Kelly P, Doherty A, Mizdrak A, Marshall S, Kerr J, Legge A, et al. High group level validity but high random error of a self-report travel diary, as assessed by wearable cameras. J Transp Heal [Internet]. 2014;1(3):190–201. Available from: http://linkinghub.elsevier.com/retrieve/pii/S2214140514000358

60. Kelly P, Doherty AR, Hamilton A, Matthews A, Batterham AM, Nelson M, et al. Evaluating the Feasibility of Measuring Travel to School Using a Wearable Camera. Am J Prev Med [Internet]. 2012 Nov;43(5):546–50. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0749379712005314

61. Oliver M, Doherty AR, Kelly P, Badland HM, Mavoa S, Shepherd J, et al. Utility of passive photography to objectively audit built environment features of active transport journeys: an observational study. Int J Health Geogr.

2013;12(1):20.

62. Nebeker C, Lagare T, Takemoto M, Lewars B, Crist K, Bloss CS, et al. Engaging research participants to inform the ethical conduct of mobile imaging, pervasive sensing, and location tracking research. Transl Behav Med [Internet]. 2016 Dec 29;6(4):577–86. Available from: https://academic.oup.com/tbm/article/6/4/577-586/4563223

63. Wilson G, Jones D, Schofield P, Martin DJ. Experiences of using a wearable camera to record activity, participation and health-related behaviours: Qualitative reflections of using the Sensecam. Digit Heal [Internet]. 2016 Jan 1;2:205520761668262. Available from: http://journals.sagepub.com/doi/10.1177/2055207616682628

64. Gage R, Leung W, Stanley J, Reeder A, Mackay C, Chambers T, et al. Studying third-parties and environments: New Zealand sun-safety research. Health Promot Int [Internet]. 2019 Jun 1;34(3):440–6. Available from: https://academic.oup.com/heapro/article/34/3/440/4748797

65. Watkins L, Aitken R, Gage R, Smith MB, Chambers TJ, Barr M, et al. Capturing the Commercial World of Children: The Feasibility of Wearable Cameras to Assess Marketing Exposure. J Consum Aff [Internet]. 2019 Dec 18;53(4):1396–420. Available from: https://onlinelibrary.wiley.com/doi/10.1111/joca.12234

66. Imtiaz MH, Hossain D, Senyurek VY, Belsare P, Tiffany S, Sazonov E. Wearable Egocentric Camera as a Monitoring Tool of Free-Living Cigarette Smoking: A Feasibility Study. Nicotine Tob Res [Internet]. 2020 Oct 8;22(10):1883–90. Available from: https://academic.oup.com/ntr/article/22/10/1883/5613972

67. Zhou Q, Wang D, Mhurchu CN, Gurrin C, Zhou J, Cheng Y, et al. The use of wearable cameras in assessing children's dietary intake and behaviours in China. Appetite [Internet]. 2019 Aug;139:1–7. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0195666318314569

68. Veatupu L, Puloka V, Smith M, McKerchar C, Signal L. Me'akai in Tonga: Exploring the Nature and Context of the Food Tongan Children Eat in Ha'apai Using Wearable Cameras. Int J Environ Res Public Health [Internet].

2019 May 14;16(10):1681. Available from: https://www.mdpi.com/1660-4601/16/10/1681

69. Laskaris Z, Milando C, Batterman S, Mukherjee B, Basu N, O'neill MS, et al. Derivation of Time-Activity Data Using Wearable Cameras and Measures of Personal Inhalation Exposure among Workers at an Informal Electronic-Waste Recovery Site in Ghana. Ann Work Expo Heal [Internet]. 2019 Oct 11;63(8):829–41. Available from: https://academic.oup.com/annweh/article/63/8/829/5537141

70. Uganda Bureau of Statistics, ICF. Uganda Demographic and Health Survey 2016. Kampala, Uganda and Rockville, Maryland, USA: UBOS and ICF; 2018.

71. World Health Organization. Indicators for assessing infant and young child feeding practices: Part 1: Definitions. World Health Organization (WHO); 2008.

72. Uganda Bureau of Statistics (UBOS). Time Use Survey 2017/2018 Report. Kampala, Uganda: UBOS; 2019.

73. Kennedy GL, Pedro MR, Seghieri C, Nantel G, Brouwer I. Dietary Diversity Score Is a Useful Indicator of Micronutrient Intake in Non-Breast-Feeding Filipino Children. J Nutr [Internet]. 2007;137(2):472–7. Available from: http://jn.nutrition.org/content/137/2/472

74. Steyn NP, Nel JH, Nantel G, Kennedy G, Labadarios D. Food variety and dietary diversity scores in children: Are they good indicators of dietary adequacy? 2006;9(5):644–50.

75. Hatløy A, Torheim LE, Oshaug A. Food variety—a good indicator of nutritional adequacy of the diet? A case study from an urban area in Mali, West Africa. Eur J Clin Nutr [Internet]. 1998 Dec 17;52(12):891–8. Available from: http://www.nature.com/articles/1600662

76. Moursi MM, Arimond M, Dewey KG, Tre S, Ruel MT, Delpeuch F, et al. Dietary Diversity Is a Good Predictor of the Micronutrient Density of the Diet of 6- to 23-Month-Old Children in Madagascar. J Nutr [Internet]. 2008;138(July):0–5. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/19022971

77. Rah JH, Akhter N, Semba RD, Pee S de, Bloem MW, Campbell AA, et al. Low dietary diversity is a predictor of child stunting in rural Bangladesh. Eur J Clin Nutr [Internet]. 2010;64(12):1393–8. Available from: http://www.nature.com/doifinder/10.1038/ejcn.2010.171

78. Arimond M, Ruel MT. Dietary Diversity Is Associated with Child Nutritional Status: Evidence from 11 Demographic and Health Surveys. J Nutr. 2004;134(10):2579–85.

79. Blackden CM, Wodon Q. Gender, Time Use, and Poverty in Sub-Saharan Africa [Internet]. World Bank Working Papers. Washington, D.C.: The World Bank; 2006. 172 p. Available from: http://elibrary.worldbank.org/doi/book/10.1596/978-0-8213-6561-8

80. Popkin BM. Time allocation of the mother and child nutrition. Ecol Food Nutr. 1980;9(1):1–14.

81. Gryboski KL. Maternal and Non-maternal Time Allocation to Infant Care, and Care during Infant Illness in Rural Java, Indonesia. Soc Sci Med. 1996;43(2):209–19.

82. Komatsu H, Malapit HJL, Theis S. Does women’s time in domestic work and agriculture affect women’s and children’s dietary diversity? Evidence from Bangladesh, Nepal, Cambodia, Ghana, and Mozambique. Food Policy. 2018;

83. Hawkes K, O’Connell JF, Blurton Jones NG. Hadza Women’s Time Allocation, Offspring Provisioning, and the Evolution of Long Postmenopausal Life Spans. Curr Anthropol [Internet]. 1997;38(4):551–77. Available from: http://www.journals.uchicago.edu/doi/10.1086/204646

84. Nordang S, Shoo T, Holmboe-Ottesen G, Kinabo J, Wandel M. Women’s work in farming, child feeding practices and nutritional status among under-five children in rural Rukwa, Tanzania. Br J Nutr. 2015;114(10):1594–603.

85. van den Bold M, Bliznashka L, Ramani G, Olney D, Quisumbing A, Pedehombga A, et al. Nutrition-sensitive agriculture programme impacts on time use and associations with nutrition outcomes. Matern Child Nutr [Internet]. 2021 Apr 10;17(2):16. Available from:

https://doi.org/10.1111/mcn.13104

86. Bamji MS, Thimayamma BVSS. Impact of women's work on maternal and child nutrition. Ecol Food Nutr [Internet]. 2000;39(1):13–31. Available from: http://www.tandfonline.com/doi/abs/10.1080/03670244.2000.9991602

87. Hirway I, Jose S. Understanding Women's Work Using Time-Use Statistics: The Case of India. Fem Econ [Internet]. 2011 Oct;17(4):67–92. Available from: http://www.tandfonline.com/doi/abs/10.1080/13545701.2011.622289

88. Fedick CB, Pacholok S, Gauthier AH. Methodological issues in the estimation of parental time – Analysis of measures in a Canadian time-use survey. Int J Time Use Res. 2005;2(1):67–87.

89. Buvinić M, King EM. Invisible No More? A Methodology and Policy Review of How Time Use Surveys Measure Unpaid Work [Internet]. United Nations Foundation, Data2X; 2018. Available from: https://data2x.org/resource-center/invisible-no-more-a-methodology-and-policy-review-of-how-time-use-surveys-measure-unpaid-work-volume-1/

90. Walthery P, Gershuny J. Improving Stylised Working Time Estimates with Time Diary Data: A Multi Study Assessment for the UK. Soc Indic Res [Internet]. 2019 Aug 15;144(3):1303–21. Available from: http://link.springer.com/10.1007/s11205-019-02074-3

91. Kanwar P, Yadav D, Sharma N. Time use pattern of hill farm women: A study in Himachal Pradesh. Himachal J Agric Res. 2003;29(1&2):84–8.

92. Szalai A. Women's time: Women in the light of contemporary time-budget research. Futures [Internet]. 1975 Oct;7(5):385–99. Available from: https://linkinghub.elsevier.com/retrieve/pii/0016328775900178

93. Engle P, Lhotská L, Armstrong H, UNICEF. The Care Initiative: Assessment, Analysis and Action to Improve Care for Nutrition [Internet]. UNICEF. New York: UNICEF; 1997. Available from: http://www.sidalc.net/cgi-bin/wxis.exe/?IsisScript=BINCA1.xis&method=post&formato=2&cantidad=1&expresion=mfn=001084

94. Floro MS, King EM. The present and future of time-use analysis in developing countries. Asia-Pacific Popul J [Internet]. 2016 Oct 24;31(1):5–42.

Available from: https://www.un-ilibrary.org/population-and-demography/the-present-and-future-of-time-use-analysis-in-developing-countries_748e616d-en

95. Johnston D, Stevano S, Malapit HJL, Hull E, Kadiyala S. Agriculture, gendered time use, and nutritional outcomes: A systematic review. Washington, D.C.: The International Food Policy Research Institute (IFPRI); 2015. (IFPRI Discussion Paper). Report No.: 01456.

96. Mcguire JS, Popkin BM. Helping Women Improve Nutrition in the Developing World. Washington, D.C.: The International Bank for Reconstruction and Development/THE WORLD BANK; 1990.

97. Floro MS. Economic restructuring, gender and the allocation of time. World Dev [Internet]. 1995 Nov;23(11):1913–29. Available from: https://linkinghub.elsevier.com/retrieve/pii/0305750X9500092Q

98. Nti CA, Inkumsah D, Fleischer G. Influence of women’s workload on their nutritional status in selected communities in Ghana. J Consum Stud Home Econ [Internet]. 1999;23(3):165–70. Available from: http://onlinelibrary.wiley.com/doi/10.1046/j.1365-2737.1999.00106.x/abstract

99. Leslie J. Women’s time: A factor in the use of child survival technologies? Health Policy Plan [Internet]. 1989;4(1):1–16. Available from: https://academic.oup.com/heapol/article-lookup/doi/10.1093/heapol/4.1.1

100. Leslie J. Women’s Work and Child Nutrition in the Third World. In: Women, Work, and Child Welfare in the Third World. 1989.

101. FAO. Minimum dietary diversity for women: An updated guide for measurement. Rome: Food and Agriculture Organization of the United Nations (FAO); 2021.

102. Grosh M, Glewwe P, editors. Designing household survey questionnaires for developing countries: Lessons from 15 years of the Living Standards Measurement Study. Vol. 3. Washington, D.C.: The World Bank; 2000. Report No.: 20731.

103. Alkire S, Meinzen-Dick R, Peterman A, Quisumbing AR, Seymour G. The women’s empowerment in agriculture index [Internet]. Oxford: Oxford

Poverty & Human Development Initiative; 2013. (OPHI WORKING PAPER). Report No.: 58. Available from: http://www.ophi.org.uk/wp-content/uploads/ophi-wp-58.pdf?cda6c1

104. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol [Internet]. 2018 Dec 19;18(1):143. Available from: https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-018-0611-x

105. Pearson A, Bottomley R, Chambers T, Thornton L, Stanley J, Smith M, et al. Measuring Blue Space Visibility and 'Blue Recreation' in the Everyday Lives of Children in a Capital City. Int J Environ Res Public Health [Internet]. 2017 May 26;14(6):563. Available from: http://www.mdpi.com/1660-4601/14/6/563

106. Kelly P, Marshall SJ, Badland H, Kerr J, Oliver M, Doherty AR, et al. An Ethical Framework for Automated, Wearable Cameras in Health Behavior Research. Am J Prev Med [Internet]. 2013;44(3):314–9. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0749379712008628

107. Hanley-Cook GT, Tung JYA, Sattamini IF, Marinda PA, Thong K, Zerfu D, et al. Minimum Dietary Diversity for Women of Reproductive Age (MDD-W) Data Collection: Validity of the List-Based and Open Recall Methods as Compared to Weighed Food Record. Nutrients [Internet]. 2020 Jul 9;12(7):2039. Available from: https://www.mdpi.com/2072-6643/12/7/2039

108. European Union. Harmonised European Time Use Surveys (HETUS) 2018 Guidelines. Luxembourg: Publications Office of the European Union; 2019.

109. Bland JM, Altman DG. Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement. Lancet. 1986 Feb;327(8476):307–10.

110. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol [Internet]. 2010 Jul;63(7):737–45. Available from:

https://linkinghub.elsevier.com/retrieve/pii/S0895435610000909

111. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine: a practical guide. Practical guides to biostatistics and epidemiology. Cambridge ; New York: Cambridge University Press; 2011. 338 p.

112. FAO and FHI 360. Minimum Dietary Diversity for Women: A Guide for Measurement. Rome: FAO; 2016.

113. World Health Organization. Indicators for assessing infant and young child feeding practices: Part 2: Measurement. Geneva: World Health Organization (WHO); 2010.

114. Ounpuu S. Seasonality, child nutrition and women's activity patterns: A case study in Chilunga Village, Malawi. [Guelph, Ontario, Canada]: University of Guelph; 1988.

115. Giavarina D. Understanding Bland Altman analysis. Biochem Medica [Internet]. 2015;25(2):141–51. Available from: http://www.biochemia-medica.com/en/journal/25/2/10.11613/BM.2015.015

116. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res. 1999;(8):135–60.

117. Gibson RS. Principles of nutritional assessment. 2nd ed. New York: Oxford University Press; 2005. 908 p.

118. Lesotho B of S. Household Budget Survey Report, Lesotho. Maseru, Lesotho; 2003.

119. Ericsson KA, Simon HA. Verbal reports as data. Psychol Rev [Internet]. 1980;87(3):215–51. Available from: http://content.apa.org/journals/rev/87/3/215

120. Willis GB. Cognitive Interviewing: A "How To" Guide. Research Triangle Park: Research Triangle Institute; 1999.

121. World Health Organization. The WHO Child Growth Standards. Geneva: World Health Organization (WHO); 2006.

122. Schreiner M. Uganda 2012 Poverty Probability Index (PPI) [Internet]. New Haven, Connecticut: Innovations for Poverty Action (IPA); 2016. Available

from: www.povertyindex.org/country/uganda

123. Yves Martin-Prevel, Arimond M, Allemand P, Wiesmann D, Ballard TJ, Deitchler M, et al. Development of a Dichotomous Indicator for Population-Level Assessment of Dietary Diversity in Women of Reproductive Age. Curr Dev Nutr [Internet]. 2017;1(12):1–10. Available from: https://academic.oup.com/cdn/article/4772173

124. United Nations Statistics Division. International Classification of Activities for Time-Use Statistics 2016 (ICATUS 2016). New York: United Nations; 2019.

125. Lombard MJ, Steyn NP, Charlton KE, Senekal M. Application and interpretation of multiple statistical tests to evaluate validity of dietary intake assessment methods. Nutr J [Internet]. 2015 Dec 22;14(1):40. Available from: http://nutritionj.biomedcentral.com/articles/10.1186/s12937-015-0027-y

126. Altman DG, Bland JM. Measurement in Medicine: The Analysis of Method Comparison Studies. Stat. 1983;(32):307–17.

127. Ji Y, Plourde H, Bouzo V, Kilgour RD, Cohen TR. Validity and Usability of a Smartphone Image-Based Dietary Assessment App Compared to 3-Day Food Diaries in Assessing Dietary Intake Among Canadian Adults: Randomized Controlled Trial. JMIR mHealth uHealth [Internet]. 2020 Sep 9;8(9):e16953. Available from: https://mhealth.jmir.org/2020/9/e16953

128. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics [Internet]. 1977;33(1):159. Available from: https://www.jstor.org/stable/2529310?origin=crossref

129. Farooq M, Doulah A, Parton J, McCrory M, Higgins J, Sazonov E. Validation of Sensor-Based Food Intake Detection by Multicamera Video Observation in an Unconstrained Environment. Nutrients [Internet]. 2019;11(3):609. Available from: https://www.mdpi.com/2072-6643/11/3/609

130. Peterson RA. A Meta-Analysis of Cronbach's Coefficient Alpha. J Consum Res. 1994;21(2):381–391.

131. Bulungu ALS, Palla L, Priebe J, Forsythe L, Katic P, Varley G, et al. Validation of a life-logging wearable camera method and the 24-h diet recall

method for assessing maternal and child dietary diversity. Br J Nutr [Internet]. 2021 Jun 14;125(11):1299–309. Available from: https://www.cambridge.org/core/product/identifier/S0007114520003530/type/journal_article

132. Gemming L. Image-assisted dietary assessment: Evaluating the potential of wearable cameras to enhance self -report in the 24-hour dietary recall method. [Auckland, New Zealand]: The University of Auckland; 2015.

133. Gemming L, Ni Mhurchu C. Dietary under-reporting: What foods and which meals are typically under-reported? Eur J Clin Nutr [Internet]. 2016;70(5):640–1. Available from: http://www.nature.com/doifinder/10.1038/ejcn.2015.204

134. Chan V, Davies A, Wellard-Cole L, Lu S, Ng H, Tsoi L, et al. Using Wearable Cameras to Assess Foods and Beverages Omitted in 24 Hour Dietary Recalls and a Text Entry Food Record App. Nutrients [Internet]. 2021 May 26;13(6):1806. Available from: https://www.mdpi.com/2072-6643/13/6/1806

135. Bruun A, Stentoft ML. Lifelogging in the Wild: Participant Experiences of Using Lifelogging as a Research Tool. In: D. Lamas, F. Loizides, L. Nacke, H. Petrie, M. Winckler & PZ, editor. Proceedings of the 17th IFIP TC13 International Conference on Human-Computer Interaction – INTERACT 2019 [Internet]. 2019. p. 431–51. Available from: http://link.springer.com/10.1007/978-3-030-29387-1_24

136. Silva AR, Pinho S, Macedo LM, Moulin CJ. Benefits of SenseCam Review on Neuropsychological Test Performance. Am J Prev Med [Internet]. 2013 Mar;44(3):302–7. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0749379712008501

137. Hodges S, Williams L, Berry E, Izadi S, Srinivasan J, Butler A, et al. SenseCam: A Retrospective Memory Aid. In 2006. p. 177–93. Available from: http://link.springer.com/10.1007/11853565_11

138. Chow TE, Rissman J. Neurocognitive mechanisms of real-world autobiographical memory retrieval: Insights from studies using wearable camera technology. Ann N Y Acad Sci [Internet]. 2017 May 26;1396(1):202–

21. Available from: https://onlinelibrary.wiley.com/doi/10.1111/nyas.13353

139. Small L, Sidora-Arcoleo K, Vaughan L, Creed-Capsel J, Chung K-Y, Stevens C. Validity and Reliability of Photographic Diet Diaries for Assessing Dietary Intake Among Young Children. ICAN Infant, Child, Adolesc Nutr [Internet]. 2009;1(1):27–36. Available from: http://journals.sagepub.com/doi/10.1177/1941406408330360

140. Wilson G, Aitken D, Hodgson P, Bailey C. The hidden impact of home adaptations: Using a wearable camera to explore lived experiences and taken-for-granted behaviours. Health Soc Care Community [Internet]. 2019 Nov 8;27(6):1469–80. Available from: https://onlinelibrary.wiley.com/doi/10.1111/hsc.12818

141. Liu W, Barr M, Pearson AL, Chambers T, Pfeiffer KA, Smith M, et al. Space-time analysis of unhealthy food advertising: New Zealand children's exposure and health policy options. Health Promot Int. 2019;1–9.

142. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet [Internet]. 2007 Oct;370(9596):1453–7. Available from: https://linkinghub.elsevier.com/retrieve/pii/S014067360761602X

143. Bulungu ALS, Palla L, Priebe J, Forsythe L, Katic P, Varley G, et al. Validation of an Automated Wearable Camera-Based Image-Assisted Recall Method and the 24-Hour Recall Method for Assessing Women's Time Allocation in a Nutritionally Vulnerable Population: The Case of Rural Uganda. Nutrients [Internet]. 2022 Apr 27;14(9):1833. Available from: https://www.mdpi.com/2072-6643/14/9/1833

144. Bulungu ALS, Palla L, Priebe J, Forsythe L, Katic P, Varley G, et al. UNDER REVIEW: Validation of an automated wearable camera-based image-assisted recall method and the 24-hour activity recall method for assessing women's time allocation in rural Uganda. Soc Indic Res. 2022;

145. Doherty AR, Moulin CJA, Smeaton AF. Automatically assisting human memory: A SenseCam browser. Memory [Internet]. 2011 Oct;19(7):785–95. Available from:

http://www.tandfonline.com/doi/abs/10.1080/09658211.2010.509732

146. Zuppinger C, Taffé P, Burger G, Badran-Amstutz W, Niemi T, Cornuz C, et al. Performance of the Digital Dietary Assessment Tool MyFoodRepo. Nutrients [Internet]. 2022 Feb 1;14(3):635. Available from: https://www.mdpi.com/2072-6643/14/3/635

147. Jia W, Wu Z, Ren Y, Cao S, Mao Z-H, Sun M. Estimating Dining Plate Size From an Egocentric Image Sequence Without a Fiducial Marker. Front Nutr [Internet]. 2021 Jan 14;7. Available from: https://www.frontiersin.org/articles/10.3389/fnut.2020.519444/full

148. Mohanty SP, Singhal G, Scuccimarra EA, Kebaili D, Héritier H, Boulanger V, et al. The Food Recognition Benchmark: Using Deep Learning to Recognize Food in Images. Front Nutr [Internet]. 2022 May 6;9. Available from: https://www.frontiersin.org/articles/10.3389/fnut.2022.875143/full

149. Gao A, Lo FP-W, Lo B. Food volume estimation for quantifying dietary intake with a wearable camera. In: Body Sensor Networks Conference 2018. IEEE; 2017.

150. Yue Y, Jia W, Fernstrom JD, Sclabassi RJ, Fernstrom MH, Yao N, et al. Food volume estimation using a circular reference in image-based dietary studies. In: Proceedings of the 2010 IEEE 36th Annual Northeast Bioengineering Conference (NEBEC) [Internet]. IEEE; 2010. p. 1–2. Available from: http://ieeexplore.ieee.org/document/5458155/

151. Sun M, Fernstrom JD, Jia W, Hackworth SA, Yao N, Li Y, et al. A Wearable Electronic System for Objective Dietary Assessment. J Am Diet Assoc [Internet]. 2010;110(1):45–7. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0002822309016897

152. Jia W, Li Y, Qu R, Baranowski T, Burke LE, Zhang H, et al. Automatic food detection in egocentric images using artificial intelligence technology. Public Health Nutr [Internet]. 2019 Mar 26;1–12. Available from: https://www.cambridge.org/core/product/identifier/S1368980018000538/type/journal_article

153. Doherty AR, Caprani N, Conaire CÓ, Kalnikaite V, Gurrin C, Smeaton AF, et

al. Passively recognising human activities through lifelogging. Comput Human Behav [Internet]. 2011 Sep;27(5):1948–58. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0747563211000963

154. Doherty AR, Smeaton AF. Automatically Segmenting LifeLog Data into Events. In: 2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services [Internet]. IEEE; 2008. p. 20–3. Available from: http://ieeexplore.ieee.org/document/4556872/

155. Talavera E, Leyva-Vallina M, Sarker MK, Puig D, Petkov N, Radeva P. Hierarchical approach to classify food scenes in egocentric photo-streams. 2018.

156. The World Bank. World Bank Country and Lending Groups [Internet]. [cited 2022 Jun 7]. Available from: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups