



# Just you wait... and fill out this survey. Discussion of the methodological aspects of waiting room surveys

Yfke P. Ongena<sup>1</sup> · Marieke Haan<sup>2</sup>

Received: 26 October 2021 / Revised: 9 February 2022 / Accepted: 8 March 2022 /  
Published online: 26 March 2022  
© The Author(s) 2022

## Abstract

A method commonly used in health care research is the waiting room (WR) survey. While patients are waiting for their appointment, they are asked to complete a questionnaire measuring their attitudes, behaviors and other characteristics. In this paper we synthesize practical guidelines for WR surveys by comparing the method with two similar approaches: public intercept (PI) surveys and drop-off-pick-up (DOPU) surveys. In this comparison we use the Total Survey Error approach Groves (Survey Methodology, Wiley, New York, 2004); (Groves in Public Opinion Quarterly 74(5): 849-879, 2010) and apply it to three case examples in which one of the three surveys is used. We take into account measurement (validity, measurement- and processing error) and representation (coverage-, sampling- and nonresponse error). From our review, we conclude that waiting room surveys, though limited to patients and their caregivers, can provide useful information on patients' perspective on health care. Response rates in waiting rooms are usually high, but often not even reported. We recommend adjustment for sampling bias by taking into account the number of visits to the hospital per respondent and sample times proportionate to the number of sample members expected on a particular time. These surveys also allow for collection of para-data; i.e., relevant information in the circumstances of a request to participate in survey research, and behavior of surveyors can easily be controlled, or investigated in an experimental design.

**Keywords** Waiting room surveys · Survey research methods · Patient attitudes · Questionnaires

## 1 Introduction

Waiting room (WR) surveys are conducted in the waiting room of a health care facility before or after clinical encounters (Hogg et al. 2010). These kinds of surveys are commonly used to study patients' attitudes, knowledge, behaviors related to consultations and

---

✉ Yfke P. Ongena  
y.p.ongena@rug.nl

<sup>1</sup> Center for Language and Cognition, University of Groningen, Groningen, The Netherlands

<sup>2</sup> Department of Sociology, University of Groningen, Groningen, The Netherlands

other characteristics. Though, in many studies, the methodological choices for WR surveys are not explained, and also response rates are not always mentioned (Pirotta et al. 2002). Guiding literature on WR surveys for practitioners is scarce and outdated (Pirotta et al. 2002; Hogg et al. 2010). Taking the Total Survey Error approach (Groves 2004; Groves and Lyberg 2010), in this paper we synthesize practical guidelines from prior studies, and our own experience in using WR surveys.

To get a better understanding of the advantages and disadvantages of WR surveys, we will compare these surveys with two other kind of surveys: (1) drop-off-pick-up surveys (DOPU; Trentelman et al. 2016) and (2) public intercept surveys (PI; Cowan 1989). In DOPU surveys, surveyors hand deliver a self-administered paper-and-pencil questionnaire at the sampling members' homes. In a WR survey, patients waiting for their appointment in the hospital or health care centre are requested to self-complete a questionnaire on particular health care topics; thus, in terms of data collection it seems similar to DOPU surveys.

A PI survey entails selecting a public place such as a shopping mall, a public park, an airport or a gas station, and approaching people who happen to be present. In general, public places provide useful opportunities for data collection for research topics with a place-based focus (Jackson-Smith et al. 2016; Rookey et al. 2012). WR surveys therefore resemble PI surveys in terms of sampling; the sampling takes place in a public environment and is based on convenience. These kinds of convenience sampling methods have been criticized (Cowan 1989; Jackson-Smith et al. 2016), but there are also beneficial factors.

In our comparison of WR versus DOPU and PI surveys we will take both the criticism and the beneficial factors into account by discussing the Total Survey Error approach (Groves et al. 2004; Groves and Lyberg 2010), while also discussing three examples of type of survey (Table 1). These examples (Due et al. 2018; Lin et al. 2014 and Lee et al. 2019) were selected based on their similar populations (elderly people), and have similar survey topics (social relations, loneliness, and/or accessibility).

## 2 A comparison of WR versus DOPU and PI surveys

The total survey error approach (Groves et al. 2004; Groves and Lyberg 2010) is a helpful framework in comparing different survey designs for statistical error properties. The model makes a distinction between error sources related to *measurement* (validity, measurement- and processing error) and *representation* (coverage-, sampling- and nonresponse error). Since most criticism on WR surveys exists with respect to errors sources related to representation (e.g., the sampling method), we will discuss this component first. Table 2 summarizes our comparison between three survey methods of WR surveys, DOPU surveys and PI surveys, and the comparison is explained in more detail in the next subsections.

### 2.1 Representation: coverage error in WR versus DOPU and PI

Coverage error refers to a mismatch between the sampling frame and the target population. DOPU surveys allow for using sampling frames that enable random sampling methods like multi-stage sampling with field listing or address-based sampling using residential address lists (Kalton et al. 2014). Thus, a target population in DOPU surveys can include the general population, and topics for data collection are not restricted. Administration of intricate survey designs, implementation of complex eligibility criteria for respondents, and introducing respondents to low interest topics can be a reason

**Table 1** Examples of WR, PI, and DOPU surveys

	Waiting room (Due et al. (2018))	Public intercept (Lin et al. 2014)	Drop-off-Pick-up (Lee et al. 2019)
Main Aim	Study the correlation between patients' descriptions of their social relations and feelings of loneliness, and their general practitioners' assessments of these	Study factors affecting train station accessibility for elderly train station users	Study the effects of neighbourhood walkability and third places on older adults' social connectedness
Population	Patients aged 65 and older in 12 general practices in Denmark and their general practitioners	Train station users aged 60 years or older, at seven highly dispersed train stations in Perth (Australia)	Homeowners aged 65 years and older living in their own single-family homes in College Station and Bryan (Texas, USA)
Sampling	During a three-week period in each practice patients were recruited; the general practitioner (GP) filled out a matching questionnaire	On two consecutive days, train users were recruited	Simple random sample of the recipient list of "over sixty-five homestead tax exemption" was obtained from the county tax office
Data Collection	Data were collected in GP waiting rooms from February to September 2014 (n = 476 completed the surveys)	Data were collected via face-to-face surveys on 7 train stations on July 31 and August 1, 2012 (n = 940 completed surveys, of which n = 122 elderly users)	Data were collected door-to-door between 9:00 a.m. and 6:00 p.m. via a drop off / pick up survey from late March to May in 2014 (n = 305 completed surveys)
Surveyors	GP staff (GP's and their staff were instructed by the first author)	No details available	'One surveyor' (no other details available)

**Table 2** Comparison of data collection characteristics of WR, DOPU and PI surveys

TSE component	Waiting room	Public intercept	Drop-off-Pick-up
Coverage error	Non-household convenience samples, allows for place-based topics	Non-household convenience samples, allows for place-based topics	Household-based sampling, unlimited topics (also sensitive) topics
Sampling error	Frequency bias and length bias, reduced if sampling with probabilities proportionate to sample size	Frequency bias and length bias, reduced if sampling with probabilities proportionate to sample size	Reduced by random sampling
Non-response: possibilities of eligibility registration	Good due to limited number of access points	Limited due to usually higher number of access points	Depends on contact-rate
Non-response: possibility of advance notification	When patients make their appointment or announced on bulletin boards	Survey can be announced in local newspaper or billboard or online advertising	Advance letters can be sent, and surveyors can notify respondents via phone
Non-response: social influence on participation	Social pressure from observing others participating	Avoiders with aversion to intercept-surveys can avoid eye-contact and walk away	Depends on persuasion from surveyor
Non-response: staff requirements	Use of existing reception staff may reduce response rate, assistants can be hired	Surveyors must be hired to approach and interview respondents	Surveyors must be hired to deliver questionnaires
Non-response: time constraints in completion	Available completion time depends on time while waiting for appointment	Immediate participation is required, interrupting one's actions	Survey can be completed within agreed time frame of usually 24 to 72 h
Non-response: information on non-respondents	Observation, or asking a small set of questions orally, or use of health records	Observation or asking a small set of questions orally	Depends on contact, on neighborhood-level observation is possible
Validity	Short questionnaire, easy combination with other measurements	Short questionnaire, easy combination with observation	Longer questionnaires are possible
Measurement error: possibilities of mode for data collection	Paper-and-pencil survey, tablets or own device via QR-code	Face-to-face interview, paper-and-pencil survey, or own device via QR-code	Paper-and-pencil survey, or own device via QR-code
Processing error	More possibilities of computer-administered surveys: less data entry errors	Fewer possibilities of computer-administered surveys: more data entry errors	Fewer possibilities of computer-administered surveys: more data entry errors

to choose a DOPU design rather than a traditional mail survey (Trentelman et al. 2016). In the DOPU example case (Lee et al. 2019), homeowners were the target group, which made their sampling procedure using addresses from the homestead tax exemption recipient list very useful. In WR surveys and PI surveys, the sampling frame is limited to people who happen to be present at the facility during data collection. Thus, sampling is based on selection of places and periods, though stratified samples of places can be implemented (Bator et al. 2011). In the WR example case (Due et al. 2018) patients were the target group, while the aim was to relate their responses to a matching questionnaire filled out by the general practitioner. Because of that match, and both persons filling out the questionnaire at the same time, recruiting participants in 12 different general practices is a useful choice. In the PI example case (Lin et al. 2014) elderly train users were the target group. Although detailed procedures are not given in their paper, we assume that all train users were approached regardless of their perceived aged, since from the original sample of 940 train users, only 122 were elderly users. Thus, a target population in WR and PI surveys does not include the general population, and topics for data collection are restricted to place-based topics.

## 2.2 Representation: sampling error in WR versus DOPU and PI

Sampling error refers to the error that occurs due to the fact that not an entire population of interest is studied. Sampling error can be reduced by drawing a random sample as is commonly done in DOPU surveys. Also, the sample size could be increased, however, risks of selection bias then still remain (Taherdoost 2017). The DOPU example case (Lee et al. 2019) indeed used simple random sampling with sample size determined by applying the probability sampling formula for a simple random sample (with a confidence level 95%, margin of error = 5% applied to the population size). In PI and WR surveys, the sample size has more impact on the length of the field period and also depends on the number of facilities included, and obviously, in waiting rooms selection of times and days of the week is restricted to the opening hours of the facility. The WR example case (Due et al., 2018) used a three-week period and 12 facilities, the times and days of the week were not mentioned in the paper. Based on statistics available on the average frequency of GP visits in Denmark, the authors argue that their target group were likely to be frequent GP visitors. Such information is useful in estimating sampling error.

Problems in PI surveys may arise in terms of frequency bias and length bias: individuals who visit a public space more frequently or who spent more time at such a facility are more likely to be oversampled (Nowell and Stanley 1991). Sampling bias may be adjusted by using frequency of visits to a facility in sample weights. Length bias in WR surveys may be affected by unpunctual patients, no-shows, walk-ins, and/or emergencies. These factors may be distributed unevenly over time, but patients' unpunctuality is generally assumed to be independent of their scheduled appointment times (Cayirli and Veral 2003). Both in WR and PI surveys, time-based sampling can be helpful in reducing frequency and length bias in sampling (Bruwer et al. 1996). A simple, but not very efficient design is to sample all eligible time periods with equal probabilities (Sudman 1980). Sampling time periods with probabilities proportionate to the number of sample members expected in the time period is more efficient (Bruwer et al. 1996). The PI example case (Lin et al. 2014) sampled during only two days, between 6 am and 6 pm, and the authors acknowledge the possibility of sampling error in their conclusion.

## 2.3 Representation: nonresponse error in WR versus DOPU and PI

Non response error refers to biases that occur due to the fact that not all sample members agree to participate in the survey. As Eckman and Kreuter (2017) note, there is often a connection between non-response and coverage error. For instance, cases may inadvertently be registered as ineligible cases while in reality they should have been considered as nonresponse. These inadequate registrations may occur as a result of sample members finding out eligibility criteria (i.e., when hearing participants need to be of certain age, a sample member may report an age outside that range). Especially in surveys where eligible sample members can observe other potential respondents being approached, they are likely to find out about the eligibility criteria and adapt their responses to selection questions accordingly. Thus, to calculate response rates adequately, it is important to correctly assess eligibility of sample members. In DOPU surveys, delivery of surveys and enumerating the number of eligible respondents depends on actual contact with sample members, and therefore often a distinction is made between response rate (including non-contacts) and cooperation rate (excluding non-contacts). Although in the DOPU example case (Lee et al. 2019) no response rate was mentioned, it can be derived from the number of usable questionnaire and approached addresses given (27.5%). In PI surveys, lack of control in access points complicates tracking response rates (i.e., the proportion of actual respondents among eligible sample members). The PI example case (Lin et al. 2014) demonstrates this difficulty, as it does not provide a response rate, and does not discuss assessment of eligibility of respondents. For WR surveys, registration of eligible respondents is more convenient since they allow for much more control over observation of patients entering and exiting the waiting room. In the WR example case (Due et al. 2018), eligibility of respondents could be assessed in order to compute an adequate response rate (62%).

### 2.3.1 Possibility of advance notification

Response rates can be increased when respondents are notified about the survey some time before the request to participate, especially when this notification is personalized (Lynn 2016; Goldstein and Jennings 2002; Dillman et al. 2014; Vogl et al. 2019). Advance (and also personalized) notification of the survey is possible in WR surveys when respondents make their appointment, either by phone, mail or email. Likewise, in DOPU surveys it is possible to send personalized advance notice letters, whereas PI surveys cannot be personally announced, but possibilities exist of announcements in local newspapers or billboard advertising. Advance notification was not mentioned in any of our three example cases.

### 2.3.2 Social influence on participation

In WR surveys and PI surveys multiple sample members are usually present at the same time. This can have both beneficial and detrimental effects on response rates. On the one hand, social validation effects may increase response rates (i.e., believing others are willing to comply increases compliance, Groves et al. 1992), but on the other hand, viewing others already complying makes the request less scarce and as a result, sample members view their contribution as less useful. Ongena et al. (2021) indeed found evidence for this scarcity effect. When fewer patients were present in the waiting room relatively more patients were willing to participate in their survey. In addition, patients waiting alone accompanied

by one or more caregivers were less likely to fully complete the questionnaire, than patients waiting on their own. Reasons that Ongena et al. (2021) suggest for this effect are a possibility of reduced anxiety or increased boredom among patients waiting alone. In PI surveys, people with an aversion to PI surveys (i.e., avoiders, see Keillor & Sutton, 1993) can avoid eye-contact or walk away from surveyors as soon as they notice a PI survey is being conducted. In DOPU surveys, potential respondents are usually unaware of behaviors of other sample members because they are contacted individually.

### 2.3.3 Staff requirements

In a WR survey, costs can be considerably low when reception staff instead of specifically hired research assistants approach potential respondents for the survey. However, this reduces possibilities of persuading reluctant respondents. In any survey method where surveyors are personally approaching potential respondents, there are increased possibilities in verbally persuading reluctant respondents when staff is specifically trained (Dijkstra and Smit 2002; Ongena and Haan 2016), which can increase response rates (Bowling 2005). Nonetheless, surveyors may differ in their impact on the representativeness of the achieved sample (e.g., Blom et al. 2011; Jäckle et al. 2013; Durrant et al. 2010). Due to the fact that WR surveys are relatively small-scale, experiments on surveyor behavior can easily be conducted (see Ongena et al. 2021). As the WR example case (Due et al. 2018) shows, staff at the facility can be instructed to recruit participants and distribute questionnaires, which is very commonly done in WR surveys. However, Due et al. 2018 do note that it was difficult to recruit more than 12 practices for participation, though the GP practices that did participate appeared to be representative in terms of geographical spread and solo versus partnership practices. In DOPU and PI surveys, staff needs to be hired to visit respondents at their homes or at public places. In PI surveys, the staff usually also administers the survey (i.e., face-to-face), which requires more time and hence increases costs. Both the PI (Lin et al. 2014) and DOPU (Lee et al. 2019) example cases provide no information on their surveyors.

### 2.3.4 Time constraints in completion

Time constraints are largest in PI surveys, where usually immediate participation is required, interrupting the respondents' current activities. In WR surveys, the survey is mostly aiming for preprocess administration (i.e., from the time of arrival to when the patient is called for their appointment, Becker & Douglass 2008) though postprocess (i.e., after the medical appointment) is possible as well. Thus, in WR surveys time is only constrained within waiting time, and respondents might find filling out a survey a useful task while waiting for their appointment. Though survey completion postprocess (i.e., after the appointment, before leaving the hospital) is an option, it is not recommended, as surveyors would have less control over actual completion. In WR surveys, the request is also not as a sudden interruption as is the case in PI surveys. Especially when waiting times are long the willingness to participate may be increased, though when waiting times are exceptionally long frustration may negatively influence compliance. In addition, the surveyor can monitor the process of filling out the questionnaire from some distance. In DOPU surveys, respondents have most time to complete a survey. Using a "doorknob" bag allows the surveyor to pick-up the completed questionnaire without further contact with the respondent.

Nevertheless, a pick-up time, for instance 24 h to 72 h needs to be agreed upon (Trentelman et al. 2016).

### 2.3.5 Collection of information from non-respondents

Non-response bias can be adjusted through weighting survey outcomes with characteristics of non-respondents, though this adjustment is still topic of debate (Kreuter et al. 2011). The ability of collecting information on non-respondents in DOPU surveys depends on interviewers' contact abilities, although information on for example neighborhood characteristics can also be collected when no contact is made (i.e., urbanicity level, presence of gates etc.). In the DOPU example case (Lee et al. 2019) such characteristics on neighborhoods seem not to have been registered, as they are not reported. In a WR survey and to some extent also in a PI survey, it is possible to collect characteristics like the non-respondent's gender, the number of people present in the waiting room or public area, time of survey administration and (in WR only) the length of time the patient had waited before being approached. Characteristics like age, race, insurance status, employment status, triage acuity level, chief complaint, and triage time using the electronic health record, can be asked from non-respondents who are willing to orally respond to a small number of questions from a surveyor or who agree with accessing their health records (Shaikh et al. 2013). In the WR example case (Due et al. 2018), elaborate information is available from non-respondents, on the basis of which it could be concluded that participants were younger (59% under 75) than non-participants (43% under 75), while gender was distributed equally. In the PI example case (Lin et al. 2014), characteristics of non-respondents seem not to have been recorded.

## 2.4 Measurement: validity in WR versus DOPU and PI

Measurement component validity refers to the adequacy of conceptualizations. Survey researchers are interested in concepts that can be measured by means of standardized questionnaire items, to be self-reported by respondents. In questionnaires, valid measurement often requires longer questionnaires. However, in PI and WR surveys, usually long surveys are not possible, whereas the DOPU survey could allow for longer questionnaires. However, when surveyors are distributing questionnaires, it is also possible to combine this task with observation of events and readily observable non-verbal behaviors onsite. For instance, Bator and colleagues (2011) were able to unobtrusively observe littering behavior and unbiasedly survey the same individuals in their PI survey of litterers and disposers, pairs of observers and interviewers. The PI example case (Lin et al. 2014), combined data collected by means of surveys was with data on land use, road networks, public transport networks, services, and patronage from governmental institutions. The WR example case (Due et al. 2018) also shows a useful means of data collection that entailed matching responses from patients with a matching questionnaire filled out by the general practitioner.

## 2.5 Measurement: measurement error in WR versus DOPU and PI

Measurement error refers to the errors, both systematic and random, that may arise in receiving responses to items. In general, social desirability concerns are lower for off-site surveys than onsite. Social desirability concerns therefore seem lower for DOPU than for PI and WR surveys. A direct comparison between on-site WR survey and an off-site mail



survey with paper-and-pencil questionnaires revealed that respondents aged below 45 reported higher patient satisfaction ratings on-site than off-site, but no such associations were found for respondents aged over 45 (Burroughs et al. 2005).

In addition, onsite research may especially have added value when specific events and experiences during the event are measured. As Wofinden (2003) states: “*the need to survey people as they make a journey or ‘immediately encounter an experience’ in order that their recall of behaviour is better than it would be in a remote location and they focus their attention on the experience.*”. Furthermore, conducting survey research on a location where people are performing activities that are of interest in the survey can create opportunities. For instance, a smart design of the survey procedure also allows for combining observation of actual behavior (i.e., littering behavior) with responses to a survey, as shown by Bator et al. (2011). In general, on site surveys allow survey researchers to ask people about their behaviours and thoughts as they encounter an experience on the spot. Since respondents are in the midst of the experience, this likely entails optimal recall of behaviour and experiences, though timing of the survey is essential, since respondents may also adapt their behavior as a consequence of becoming aware of the survey. From our three example cases, the PI survey (Lin et al. 2014) most clearly shows this connectedness of the experience (begin in a train station while just having commuted to that station) and the questions asked in the questionnaire (accessibility of the train station).

### 2.5.1 Possibilities of mode for data collection

Completion of the survey in WR and DOPU surveys is mostly through a paper-and-pencil questionnaire (as is true for the WR and DOPU example cases), while PI surveys are mostly completed by means of a face-to-face survey, as is true for the PI example case). Moving to electronic web surveys is an obvious strategy in this digital era. Slater and Kiran (2016) conclude that e-mail surveys offer a convenient, low-cost method of regularly surveying patients to improve quality of care, but patients living in low-income neighborhoods are likely to be underrepresented, so this change can also affect coverage rates. Due to the fact that in WR surveys respondents are seated and stay within view of the surveyor, in these surveys it is possible to use tablets for digital completion. Although use of tablets has not demonstrated to increase response rates for on-site surveys, respondents do enjoy an electronic mode and prefer this mode over the traditional paper-and-pencil survey (Davis et al. 2012; Fanning and McAuley 2014). The advantages of electronic administration over paper-and-pencil administration in terms of response rates may be greater for longer surveys; the more pages, the less likely respondents will comply. In the early days of tablets, respondents were also fascinated by the technology (Pfaffenberg et al. 2014). For a single survey, use of tablets may not be very cost effective, but for multi-year projects with a sample sizes over 100 respondents and survey lengths over 5 pages, tablets are expected to be more cost effective and efficient than paper-based surveys (Hassler et al. 2018; Leisher 2014). Tablet surveys also have lower average completion time and were found to be fully complete more frequently than questionnaires completed on paper (Hassler et al. 2018). In a mixed-device survey with a student population however, no effects in data quality and psychometric survey properties were found (Ravert et al. 2015). The expected number of daily survey participants and the time necessary to complete a questionnaire will help in estimating the required number of tablets and surveyors (Hassler et al. 2018).

However, administration with a tablet to be used by many respondents requires hygiene measures (Pfaffenberg et al. 2014) and the number of tablets available may be a limiting

factor in surveying all sample members on a given time (Hassler et al. 2018). A good alternative to tablet is using QR-codes for allowing respondents to complete the survey on their own smartphone or tablet, while in DOPU surveys it is also possible to give a simple URL and login for completion on a PC. In public transport surveys (Guirao et al. 2015) and student surveys (Snyder et al. 2018; Onimowo et al. 2020; Faggiano et al. 2021) QR codes have shown to reduce administrative time and reaching reasonable response rates (Snyder et al. 2018; Faggiano et al. 2021), and as such being an eco-friendly alternative to paper-and-pencil questionnaires (Onimowo et al. 2020). Since they only need to be scanned, they are also preferred over bit.ly short links that require users to type, with the possibility of errors (Guirao et al. 2015).

Tablets and or QR codes could also be used in a mixed-mode design allowing respondents to choose their preferred mode. Although in other mixed mode designs no positive effects for response rates were found (Haan et al. 2014), commonly highest response rates are found when respondents were offered a choice between paper-and-pencil and web modes (Börkan 2010; Greenlaw and Brown-Welty 2009; Kiernan et al. 2005; McCabe 2004; Schonlau et al. 2003; and Sax et al. 2003).

## 2.6 Measurement: processing error in WR versus DOPU and PI

After data collection, processing errors can occur when preparing data for analysis. These errors, that are often underestimated (Jedinger et al. 2018), can include data entry (for non-computer-assisted surveys), and cleaning and editing of data (i.e., coding open-ended answers, assigning variable and value labels, handling missing values, and implementation of survey weights) and tabulation of survey data (Biemer 2010). Computer-assisted data collection, that can be used most easily in WR, and only to a smaller extent in DOPU and PI surveys, can prevent a large number of errors, provided that routing errors and implausible values are taking into account when pretesting the software. An important issue in surveys administered in different facilities, is that different selection probabilities in sampling due to differences in waiting room circumstances should be corrected for in sampling weights. Decisions on characteristics used as basis for the weights, their origin and the weighting method should be described in the methods section (Jedinger et al. 2018).

## 3 Discussion

From a comparison of WR, DOPU and PI surveys, we can conclude that though coverage is limited to the target population of health care users, WR surveys yield a useful sample when assessing patients' perspective on health care. However, a justification of choices in selecting potential respondents is important, and problems in representation may arise with specific groups of patients who do not want to participate. An important advantage of WR surveys is that they do not suffer from the problem of non-contact as household surveys like DOPU surveys do. However, some level of non-contact may arise with patients who did not show up at all for their appointment. Due to the fact that non-verbal visual contact (and in most cases also verbal contact) is available, in WR surveys eligibility criteria can be clearly assessed, which allows for adequate determination of response rates. The mere fact that para-data is available; i.e., relevant information in the circumstances of a request to participate in survey research (Kreuter 2013), is helpful in non-response research. In addition, WR surveys allow for easy control of behavior of surveyors, which in turn is useful in experimental design. The

combination of control over surveyors and availability of para-data creates insightful opportunities in investigating the persuasion of respondents.

Importantly, response rates and other relevant methodological details in WR surveys are often not reported. Therefore, it is also not known whether for instance frequency and length bias are taken into account in adjusting sampling weights, or whether sampling is adjusted to sampling times on the probabilities proportionate to the number of sample members expected in the time period (Bruwer et al. 1996). For potential respondents who refuse to participate, it is still possible to note down their gender and, if respondents comply with an ultra-brief interview, questions on their age, general frequency of visits and length of stay could be asked quickly. Similarly, for respondents who give consent for the full questionnaire such questions could be asked at the very start of the questionnaire, in order to account for these characteristics in incomplete surveys, for respondents who were called to their appointment prior to finalizing answering all items.

The fact that a WR survey is conducted on a specific location where respondents have full access to their thoughts about the experience (i.e., waiting for their health care appointment) on that location could be exploited to greater extent. Not only does it create an opportunity for optimal measurement in terms of recall (though data could be compromised in terms of social desirability bias), measurement can also be combined with observation of actual behavior and physiological measurement (i.e., biological markers, that can improve the validity of survey measures (Langhaug et al. 2010; Holmes et al. 2007).

Thus, in short, we would recommend:

- Setting up a WR survey with staff trained as surveyors and observers;
- In observation take into account variables to inform about the circumstances of data-collection (i.e., number of people present, differentiated in patients and caregivers, time and date of request and number of staff present);
- In short interviews with respondents and/or in the main questionnaire, include questions with respect to the frequency of visits and type of visit (i.e., as specific as possible in terms of length of stay) and ask about the number of visits to the department to adjust for sampling bias;
- Sample times proportionate to the number of sample members expected on a particular time.

## Declarations

**Conflict of interest** The authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. Yfke P. Ongena declares that she has no conflict of interest. Marieke Haan declares that she has no conflict of interest.

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the

material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bator, R.J., Bryan, A.D., Wesley Schultz, P.: Who gives a hoot? Intercept surveys of litterers and disposers. *Environ. Behav.* **43**(3), 295–315 (2011). <https://doi.org/10.1177/2F0013916509356884>
- Becker, F., Douglass, S.: The ecology of the patient visit: physical attractiveness, waiting times, and perceived quality of care. *The Journal of Ambulatory Care Management* **31**(2), 128–141 (2008). <https://doi.org/10.1097/01.JAC.0000314703.34795.44>
- Biemer, P.P.: Total survey error: Design, implementation, and evaluation. *Public Opin. Q.* **74**(5), 817–848 (2010). <https://doi.org/10.1093/poq/nfq058>
- Blom, A.G., de Leeuw E.D., & Hox, J.J.: Interviewer effects on nonresponse in the European Social Survey. *Journal of Official Statistics*, **27**(2), 359–377 (2011). <https://www.scb.se/contentassets/ca21efb41fee47d293bbec5bf7be7fb3/interviewer-effects-on-nonresponse-in-the-european-social-survey.pdf>
- Börkan, B.: The mode effect in mixed-mode surveys: Mail and web surveys. *Soc. Sci. Comput. Rev.* **28**, 371–380 (2010). <https://doi.org/10.1177/0894439309350698>
- Bowling, A.: Mode of questionnaire administration can have serious effects on data quality. *J. Public Health* **27**(3), 281–291 (2005). <https://doi.org/10.1093/pubmed/fdi031>
- Bruwer, J.D.W., Haydam, N.E., Lin, B.: Reducing bias in shopping mall-intercept surveys: The time-based systematic sampling method. *South African Journal of Business Management* **27**(1 & 2), 9–16 (1996). <https://doi.org/10.4102/sajbm.v27i1/2.803>
- Burroughs, T. E., Waterman, B. M., Gilin, D., Adams, D., McCollegan, J., & Cira, J.: Do on-site patient satisfaction surveys bias results? *The Joint Commission Journal on Quality and Patient Safety* **31**(3), 158–166 (2005). [https://doi.org/10.1016/S1553-7250\(05\)31021-X](https://doi.org/10.1016/S1553-7250(05)31021-X)
- Cayirli, T., Veral, E.: Outpatient scheduling in health care: a review of literature. *Prod. Oper. Manag.* **12**(4), 519–549 (2003). <https://doi.org/10.1111/j.1937-5956.2003.tb00218.x>
- Cowan, C.D.: Mall intercepts and clinical trials: The philosophy of inference from different types of research designs. *Marketing Research* **1**(1), 15–22 (1989)
- Davis, S.K., Thompson, J.L., Schweizer, S.E.: Innovations in on-site survey administration: Using an iPad interface at National Wildlife Refuges and National Parks. *Hum. Dimens. Wildl.* **17**(4), 282–294 (2012). <https://doi.org/10.1080/10871209.2012.673242>
- Dillman, D.A., Smyth, J.D., & Christian, L.M.: *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. John Wiley & Sons (2014).
- Dijkstra, W. & Smit, J.: Persuading reluctant recipients in telephone surveys. In: R. Groves, D. Dillman, J. Eltinge & R. Little (Eds.), *Survey nonresponse* (pp. 121–134). New York: John Wiley & Sons (2002).
- Due, T.D., Sandholdt, H., Siersma, V.D., Waldorff, F.B.: How well do general practitioners know their elderly patients' social relations and feelings of loneliness? *BMC Fam. Pract.* **19**(1), 1–8 (2018). <https://doi.org/10.1186/s12875-018-0721-x>
- Durrant, G.B., Groves, R.M., Staetsky, L., Steele, F.: Effects of interviewer attitudes and behaviors on refusal in household surveys. *Public Opin. Q.* **74**(1), 1–36 (2010). <https://doi.org/10.1093/poq/nfp098>
- Eckman, S., & Kreuter, F.: The undercoverage-nonresponse trade-off. In P.P. Biemer, E.D. de Leeuw, S. Eckman, B. Edwards, F. Kreuter, L.E. Lyberg, C. Tucker, & B.T. West (Eds.), *Total survey error in practice* (pp. 95–113). New York: John Wiley & Sons (2017).
- Faggiano, A., Bursi, F., Santangelo, G., Tomasi, C., Sforza, C., Faggiano, P., Carugo, S.: Global Cardiovascular Risk Profile of Italian Medical Students Assessed by a QR Code Survey. Data from UNIMI HEART SURVEY: Does Studying Medicine Hurt? *J. Clin. Med.* **10**(7), 1343 (2021). <https://doi.org/10.3390/jcm10071343>
- Fanning, J., McAuley, E.: A comparison of tablet computer and paper-based surveys in healthy aging research. *J. Med. Internet Res.* **16**, 1–10 (2014). <https://doi.org/10.2196/resprot.3291>
- Goldstein, K.M., Jennings, M.K.: The effect of advance letters on cooperation in a list sample telephone survey. *Public Opin. Q.* **66**(4), 608–617 (2002). <https://doi.org/10.1086/343756>
- Greenlaw, C., Brown-Welty, S.: A comparison of web-based and paper-based survey methods: Testing assumptions of survey mode and response cost. *Eval. Rev.* **33**, 464–480 (2009). <https://doi.org/10.1177/0193841X09340214>

- Groves, R.M., Lyberg, L.: Total survey error: Past, present, and future. *Public Opin. q.* **74**(5), 849–879 (2010). <https://doi.org/10.1093/poq/nfq065>
- Groves, R.M., Fowler, F., Couper, M.P., Singer, E., & and Tourangeau, R.: *Survey Methodology*. New York: Wiley (2004).
- Groves, R.M., Cialdini, R.B., Couper, M.P.: Understanding the decision to participate in a survey. *Public Opin. Q.* **56**(4), 475–495 (1992). <https://doi.org/10.1086/269338>
- Guirao, B., García, A., López, M.E., Acha, C., Comendador, J.: New QR survey methodologies to analyze user perception of service quality in public transport: the experience of Madrid. *J. Public Transp.* **18**(3), 5 (2015). <https://doi.org/10.5038/2375-0901.18.3.5>
- Haan, M., Ongena, Y.P., Aarts, C.W.A.M.: Reaching hard-to-survey populations: Mode choice and mode preference. *J Official Statistics* **30**(2), 355–379 (2014). <https://doi.org/10.2478/jos-2014-0021>
- Hassler, K., Pearce, K.J., Serfass, T.L.: Comparing the efficacy of electronic-tablet to paper-based surveys for on-site survey administration. *Int. J. Soc. Res. Methodol.* **21**(4), 487–497 (2018). <https://doi.org/10.1080/13645579.2018.1432403>
- Hogg, W., Johnston, S., Russell, G., Dahrouge, S., Gyorfi-Dyke, E., Kristjansson, E.: Conducting waiting room surveys in practice-based primary care research: A user's guide. *Can. Fam. Physician* **56**, 1375–1376 (2010)
- Holmes, M.D., Powell, I.J., Campos, H., Stampfer, M.J., Giovannucci, E.L., Willett, W.C.: Validation of a food frequency questionnaire measurement of selected nutrients using biological markers in African-American men. *Eur. J. Clin. Nutr.* **61**(11), 1328–1336 (2007). <https://doi.org/10.1038/sj.ejcn.1602641>
- Jäckle, A., Lynn, P., Sinibaldi, J., Tipping, S.: The effect of interviewer experience, attitudes, personality and skills on respondent co-operation with face-to-face surveys. *Survey Res Methods* **7**(1), 1–15 (2013). <https://doi.org/10.18148/srm/2013.v7i1.4736>
- Jackson-Smith, D., Dolan, M., Holyoak, G., Flint, C.G., Trentelman, C.K., Thomas, B., & Ma, G.: Effectiveness of the drop-off/pick-up survey methodology in different neighborhood types. *Journal of Rural Social Sciences*, **31**(3), 35. (2016). available at: <https://egrove.olemiss.edu/cgi/viewcontent.cgi?article=1070&context=jrss>
- Jedinger, A., Watteler, O., Förster, A.: Improving the quality of survey data documentation: A total survey error perspective. *Data* **3**(4), 45 (2018). <https://doi.org/10.3390/data3040045>
- Kalton, G., Kali, J., Sigman, R.: Handling frame problems when address-based sampling is used for in-person household surveys. *Journal of Survey Statistics and Methodology* **2**, 283–304 (2014). <https://doi.org/10.1093/jssam/smu013>
- Keillor, B.D., Sutton, R.J.: An Investigation of avoidance behavior in mall-intercept market research. *J. Mark. Manag.* **3**, 50–56 (1993)
- Kiernan, N., E., Kiernan, M., Oyler, M.A., & Gilles, C.: Is a web survey as effective as a mail survey? A field experiment among computer users. *Am. J. Eval.* **26**, 245–252 (2005). <https://doi.org/10.1177/1098214005275826>
- Kreuter, F., Olson, K., Wagner, J., Yan, T., Ezzati-Rice, T.M., Casas-Cordero, C., Lemay, M., Peytchev, A., Groves, R.M., Raghunathan, T.: Using proxy measures and other correlates of survey outcomes to adjust for non-response Examples from multiple surveys. *Journal of the Royal Statistical Society, Ser B* **73**(2), 389–407 (2011). <https://doi.org/10.1111/j.1467-985X.2009.00621.x>
- Kreuter, F. (Ed.): *Improving surveys with paradata: Analytic uses of process information*. New York: Wiley (2013).
- Langhaug, L.F., Sherr, L., Cowan, F.M.: How to improve the validity of sexual behaviour reporting: systematic review of questionnaire delivery modes in developing countries. *Tropical Med. Int. Health* **15**(3), 362–381 (2010). <https://doi.org/10.1111/j.1365-3156.2009.02464.x>
- Lee, J.H., Tan, T.H.: Neighborhood walkability or third places? Determinants of social support and loneliness among older adults. *J. Plan. Educ. Res.* (2019). <https://doi.org/10.1177/0739456x19870295>
- Leisher, C.: A comparison of tablet-based and paper-based survey data collection in conservation projects. *Social Sciences* **3**, 264–271 (2014). <https://doi.org/10.3390/socsci3020264>
- Lin, T.G., Xia, J.C., Robinson, T.P., Goulias, K.G., Church, R.L., Olaru, D., Tapin, J., Han, R.: Spatial analysis of access to and accessibility surrounding train stations: A case study of accessibility for the elderly in Perth, Western Australia. *J. Transp. Geogr.* **39**, 111–120 (2014). <https://doi.org/10.1016/j.jtrangeo.2014.06.022>
- Lynn, P.: Targeted appeals for participation in letters to panel survey members. *Public Opin. Q.* **80**(3), 771–782 (2016). <https://doi.org/10.1093/poq/nfw024>
- McCabe, S.E.: Comparison of web and mail surveys in collecting illicit drug use data: A randomized experiment. *J. Drug Educ.* **34**, 61–72 (2004)
- Nowell, C., Stanley, L.R.: Length-biased sampling in mall intercept surveys. *J. Mark. Res.* **28**(4), 475–479 (1991). <https://doi.org/10.2190/4HEY-VWXL-DVR3-HAKV>

- Ongena, Y., Haan, M.: Using interviewer-respondent interaction coding as a manipulation check on interviewer behavior in persuading CATI respondents. *Surv. Pract.* **9**(2), 1–7 (2016). <https://doi.org/10.29115/SP-2016-0010>
- Ongena, Y.P., Haan, M., Kwee, T.C., Yakar, D.: Shaking hands in a busy waiting room The effects of the surveyor's introduction and people present in the waiting room on the response rate. *Survey Methods* (2021). <https://doi.org/10.13094/SMIF-2021-00002>
- Onimowo, J.O., Knowles, G., Wrighton, G., Shah, M.: Use of quick response (QR) codes to achieve timely feedback in clinical simulation settings. *BMJ Simulation and Technology Enhanced Learning* **6**(3), 172–174 (2020)
- Pfaffenberg, C.J., Runnels, C.B., Darville, R.: The Ipad2: A survey facilitator. *J Informat Syst Technol Plan* **7**(18), 172–174 (2014). <https://doi.org/10.1136/bmjstel-2018-000426>
- Pirotta, M., Gunn, J., Harrison, D.: Measuring prevalence: Accurate sampling in general practice waiting room surveys: methodological issues. *Aust. N. Z. J. Public Health* **26**, 152–155 (2002). <https://doi.org/10.1111/j.1467-842X.2002.tb00909.x>
- Ravert, R.D., Gomez-Scott, J., Donnellan, M.B.: Equivalency of paper versus tablet computer survey data. *Educ. Res.* **44**(5), 308–310 (2015). <https://doi.org/10.3102/0013189X15592845>
- Rookey, B., D., Le, L., Littlejohn, M., & Dillman, D.A.: Understanding the resilience of mail-back survey methods An analysis of twenty years of change in response rates to National Park Surveys. *Soc. Sci. Res.* **41**, 1404–1414 (2012). <https://doi.org/10.1016/j.ssresearch.2012.06.004>
- Sax, L.J., Gillmartin, S., Bryant, A.: Assessing response rates and nonresponse bias in web and paper surveys. *Res. High. Educ.* **44**, 409–432 (2003). <https://doi.org/10.1023/A:1024232915870>
- Schonlau, M., Asch, B., Du, C.: Web surveys as part of a mixed-mode strategy for populations that cannot be contacted by E-mail. *Soc. Sci. Comput. Rev.* **21**, 218–222 (2003). <https://doi.org/10.1177/0894439303021002007>
- Shaikh, S.B., Witting, M.D., Winters, M.E., Brodeur, M.N., & Jerrard, D.A. (2013) Support for a waiting room time tracker a survey of patients waiting in an urban ED. *J. Emergency Medicine*, *44*(1): 225–229 Doi: <https://doi.org/10.1016/j.jemermed.2012.02.053>
- Slater, M., Kiran, T.: Measuring the patient experience in primary care: Comparing e-mail and waiting room survey delivery in a family health team. *Can. Fam. Physician* **62**(12), e740–e748 (2016)
- Snyder, M., Nguyen, D., Womack, J., Bunt, C., Westerfield, K., Bell, A., Ledford, C.: Testing Quick Response QR codes as an innovation to improve feedback among geographically-separated clerkship sites. *Fam. Med.* **50**(3), 188–194 (2018). <https://doi.org/10.22454/FamMed.2018.936023>
- Sudman, S.: Improving the quality of shopping center sampling. *J. Mark. Res.* **17**(4), 423–431 (1980). <https://doi.org/10.1177/002224378001700402>
- Taherdoost, H. (2017) Determining sample size; how to calculate survey sample size. *International Journal of Economics and Management Systems*
- Trentelman, C.K., Petersen, K.A., Irwin, J., Ruiz, N., & Szalay, C.S.: The case for personal interaction: Drop-off/pick-up methodology for survey research. *Journal of Rural Social Sciences*, *31*(3), 68 (2016). Available at: <https://egrove.olemiss.edu/cgi/viewcontent.cgi?article=1071&context=jrss>
- Vogl, S., Parsons, J.A., Owens, L.K., Lavrakas, P.J.: Experiments on the effects of advance letters in surveys. In: Lavrakas, P.J., Traugott, M.W., Kennedy, C., Holbrook, A.L., de Leeuw, E.L., West, B.T. (eds.) *Experimental Methods in Survey Research: Techniques that Combine Random Sampling with Random Assignment*, pp. 89–110. Wiley, New York (2019)
- Wofinden, D.: Non-Household Surveys. In P. Stopher, & P. Jones (Eds), *Transport Survey Quality and Innovation* (pp. 377–402). Oxford: Pergamon Press. (2003).