Tracking falls among elders is challenging. In this reliability study, which took place between October 2007 and February 2008, the authors compared participants' daily recordings of falls on calendars with a telephone survey of recall of falls over the previous 3 months within the population-based MOBILIZE Boston Study cohort, a cohort of 765 elders. From the cohort, 218 participants were randomly selected. Falls were tracked prospectively on daily calendars (mailed back monthly). Telephone recalls of falls over the previous 3 months were conducted in January and February 2008. Agreement, sensitivity, and specificity were calculated to compare the occurrence of falls as determined by 3-month recall with falls recorded by daily calendar (gold standard) during the same 3-month period. Results showed good agreement between recall and calendars: 27 persons reported a fall by both methods. However, while the 3-month recall correctly classified persons who did not fall (164 persons by both methods), it missed 25% of participants who fell (of 36 participants with a calendar-reported fall, 9 did not report a fall by telephone recall). Kappa was 0.74 (95% confidence interval: 0.68, 0.80), sensitivity was 75%, and specificity was 96%. Retrospective 3-month recall of falls resulted in underreporting of falls by as much as 25% compared with daily calendars. Calendars should be considered the preferred method of ascertaining falls in longitudinal studies.

Abbreviations: MMSE, Mini-Mental State Examination; MOBILIZE, Maintenance of Balance, Independent Living, Intellect, and Zest in the Elderly.

Falls are common among older persons in the United States (1). Although approximately 5%–10% of falls result in death or serious injury in this age group, it is estimated that 30%–60% of older persons living in the community suffer a fall each year (2). Most of these estimates come from longitudinal studies in which people self-report falls using simple questionnaires or daily calendars (3–7). The latter approach is considerably time-consuming and, when employed in research settings, requires adequate staffing to track calendars and verify self-reported falls with phone calls to participants. Ascertainment of falls in elders is challenging, yet to truly assess the effectiveness of a study, it is important to accurately track the number of falls.

Various methods have been suggested for tracking falls in community-dwelling elders, including retrospective recall at various intervals (e.g., weekly, bimonthly, quarterly, annually) and prospective documentation with a calendar or diary. The calendar method of falls reporting has been validated and used successfully in longitudinal studies of falls (4, 8, 9). In fact, while no one method has been declared the most effective way to track falls (10), the prospective method of using a falls calendar or diary is often considered the gold standard (5, 8–15). Specifically, calendar-based postcards have been used as a gold standard in studies evaluating recall of falls for various time periods, most often monthly or quarterly, in studies with elderly participants.
(8–10, 16, 17). The falls calendar or diary is expected to be more accurate than retrospective recall over varying time periods. However, people frequently mail incomplete calendars or do not mail completed calendars back (5, 9, 11), necessitating the need for staff to contact participants for follow-up.

Therefore, we conducted a study to compare the performance of an alternative approach of administering a single telephone survey to ascertain recall of falls over the past 3 months with falls reported using the calendar method. In this paper, we present results from a reliability study that compared falls recorded daily on a calendar postcard mailed back at the end of each month with data from a telephone recall of falls occurring over the same 3-month period that was administered in a subset of the cohort. Our aim was to determine whether retrospective recall of falls over a 3-month period could accurately track falls in comparison with the standard method of completing prospective monthly falls calendars in a population-based sample of older adults.

MATERIALS AND METHODS

MOBILIZE Boston Study participants

Participants for this study (n = 218) were a subset of participants in the MOBILIZE [Maintenance of Balance, Independent Living, Intellect, and Zest in the Elderly] Boston Study. The MOBILIZE Boston Study is a prospective, population-based cohort study of novel risk factors for falls among 765 older adults (489 women and 276 men; average age = 78 years) in the Boston, Massachusetts, area. Cohort members were enrolled from September 2005 to December 2007. Recruitment and study methods have been described in detail elsewhere (14, 15). In brief, inclusion criteria included being at least 70 years of age (or 65 years if the spouse of an enrollee), having the ability to communicate in English, and being able to walk 20 feet (6 m) without assistance from another person (walking aids were permitted). Exclusion criteria included severe vision or hearing deficits and cognitive impairment (Mini-Mental State Examination (MMSE) score < 18) (18).

For the current study, we determined that we would need 200 subjects to be able to capture sufficient numbers of falls for comparison of the 2 methods. Therefore, we randomly selected 218 participants from the MOBILIZE Boston Study cohort, with oversampling to account for possible attrition or refusal. To be selected for the current study, cohort members had to have completed the baseline assessment (where participants were trained to complete the calendars) and had to have been enrolled for at least 3 months. Selected persons were sent a postcard announcing the survey a week before study staff called them regarding the study. The institutional review board of Hebrew SeniorLife (Boston, Massachusetts) approved this study, and participants provided written and verbal informed consent.

Ascertainment of falls via falls calendar

In the MOBILIZE Boston Study, a fall is defined as any event in which any part of the body above the ankle hits the floor, the ground, or a lower surface, including falls that occur on stairs. Falls are tracked prospectively using a falls calendar. Each page of the monthly calendar is a postage-prepaid folding postcard (5) that is mailed back to the study investigators at the end of each month (see Figure 1 for an example). Participants are trained at the baseline visit to record their experience on the calendar each day, either placing an “N” in the box for the appropriate day if no fall has occurred that day or placing an “F” in the box if a fall has occurred that day. Participants are instructed to hang their falls calendars in a highly visible and frequently visited location in the home, such as the front of the refrigerator (the back of the calendar is magnetized). In the MOBILIZE Boston Study, as in other studies (5, 9, 11), up to 30% of participants require telephone calls to remind them to return their postcards monthly, to clarify unclear data entries made on the calendars, or to reiterate instructions on completing the calendars. This process is labor-intensive, resulting in a significant expenditure of resources to retrieve the missing data, but is effective in retrieving data on nearly every participant. Fewer than 1% of falls calendar data remain incomplete by the use of these methods, similarly to other field studies (2, 9, 17). Upon receipt of the monthly postcard calendars, whenever a fall is reported, a follow-up call is made to the participant to confirm the fall and collect details on the fall.

Ascertainment of falls via telephone survey of 3-month recall

A telephone survey was administered over a 6-week period, from early January 2008 to mid-February 2008. The telephone survey required approximately 5–10 minutes to administer. During this single phone call, participants were told about the new study and asked if they were willing to participate by answering a very brief set of questions over the phone. The survey assessed falls occurring in the previous 3 months and incorporated the names of those

| FEBRUARY 2008 |
|---|---|---|---|---|---|---|
| SUN | MON | TUE | WED | THURS | FRI | SAT |
| 1   | 2   | 3   | 4   | 5   | 6   | 7   |
| 8   | 9   | 10  | 11  | 12  | 13  | 14  |
| 15  | 16  | 17  | 18  | 19  | 20  | 21  |
| 22  | 23  | 24  | 25  | 26  | 27  | 28  |

**Figure 1.** Falls calendar used each month in the MOBILIZE Boston Study. Boston, Massachusetts, October 2007–February 2008. (MOBILIZE, Maintenance of Balance, Independent Living, Intellect, and Zest in the Elderly).
Interviewers conducting the telephone survey were experienced in survey administration in this study population and received specific training for the current study. In accordance with the parent study protocol, we instructed staff to make up to 6 calls to contact the participant and administer the telephone survey.

Statistical analysis

The main outcome for our study was whether a participant reported the occurrence of any fall for a 3-month time period. Baseline characteristics of the current sample were compared with those of the MOBILIZE Boston Study cohort using $t$ tests for mean measures and chi-squared tests for categorical measures. To compare the performance of the telephone survey ascertaining 3-month recall of falls with that of the falls calendar approach, we calculated chance-corrected agreement (kappa), sensitivity (proportion of calendar falls identified by survey recall), and specificity, considering the falls calendar the “gold standard.” Analyses were performed using STATA software, version 10.0 (Stata Corporation, College Station, Texas).

RESULTS

From the randomly selected sample of 218 MOBILIZE Boston cohort members, 12 persons could not be contacted during the reliability study period (October 2007–February 2008). This left 206 cohort members who participated in the reliability study. All participants who were reached agreed to participate and completed the telephone survey.

Baseline characteristics of the reliability study sample (see Table 1) were similar to those of the parent cohort. Among the 206 participants, 63% were female, and the mean age was 78.2 years (standard deviation, 5.4). The sample selected for this study had characteristics similar to those of the parent cohort, as was expected, since the subset was a random sample of nearly one-third of the parent cohort. Lower cognitive status, as indicated by MMSE scores of 18–23, was somewhat less common in the current study sample (8.7%) than in the parent MOBILIZE Boston cohort (12.0%).

Agreement between the falls calendar and the telephone survey method is presented in Table 2. Overall, there was good agreement between recall and calendars: 27 persons reported a fall by both methods. However, while the

Table 1. Characteristics of 206 Participants in a Falls Reliability Study and Participants in the MOBILIZE Boston Study Cohort*, Boston, Massachusetts, October 2007–February 2008

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Reliability Study</th>
<th>MOBILIZE Boston Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>%</td>
</tr>
<tr>
<td>Female sex</td>
<td>130</td>
<td>63.1</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>163</td>
<td>79.1</td>
</tr>
<tr>
<td>Black</td>
<td>36</td>
<td>17.5</td>
</tr>
<tr>
<td>Other</td>
<td>7</td>
<td>3.4</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>27</td>
<td>13.2</td>
</tr>
<tr>
<td>High school graduation</td>
<td>66</td>
<td>32.2</td>
</tr>
<tr>
<td>College graduation or higher</td>
<td>112</td>
<td>54.6</td>
</tr>
<tr>
<td>Body mass index</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;25</td>
<td>60</td>
<td>29.3</td>
</tr>
<tr>
<td>25–29</td>
<td>91</td>
<td>44.4</td>
</tr>
<tr>
<td>≥30</td>
<td>54</td>
<td>26.3</td>
</tr>
<tr>
<td>Mini-Mental State Examination score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–23</td>
<td>18</td>
<td>8.7</td>
</tr>
<tr>
<td>24–28</td>
<td>116</td>
<td>56.3</td>
</tr>
<tr>
<td>29–30</td>
<td>72</td>
<td>35.0</td>
</tr>
<tr>
<td>Self-rated health</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good or excellent</td>
<td>179</td>
<td>86.9</td>
</tr>
<tr>
<td>Fair or poor</td>
<td>27</td>
<td>13.1</td>
</tr>
<tr>
<td>Hearing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No difficulty</td>
<td>96</td>
<td>46.6</td>
</tr>
<tr>
<td>A little difficulty</td>
<td>39</td>
<td>18.9</td>
</tr>
<tr>
<td>Some difficulty</td>
<td>50</td>
<td>24.3</td>
</tr>
<tr>
<td>A lot of difficulty</td>
<td>21</td>
<td>10.2</td>
</tr>
<tr>
<td>No. of self-reported falls in prior year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>128</td>
<td>62.2</td>
</tr>
<tr>
<td>1</td>
<td>52</td>
<td>25.2</td>
</tr>
<tr>
<td>≥2</td>
<td>26</td>
<td>12.6</td>
</tr>
</tbody>
</table>

Abbreviation: MOBILIZE, Maintenance of Balance, Independent Living, Intellect, and Zest in the Elderly.

Table 2. Agreement* Between Falls Calendar and Telephone Survey Ascertainment of Falls, MOBILIZE Boston Study, Boston, Massachusetts, October 2007–February 2008

<table>
<thead>
<tr>
<th>Telephone Survey</th>
<th>Falls Calendar</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fall</td>
</tr>
<tr>
<td>Fall</td>
<td>27</td>
</tr>
<tr>
<td>No fall</td>
<td>9</td>
</tr>
</tbody>
</table>

Abbreviation: MOBILIZE, Maintenance of Balance, Independent Living, Intellect, and Zest in the Elderly.

* Sensitivity = 27/(27 + 9) = 75%. Specificity = 164/(164 + 6) = 96%. Percent agreement = (27 + 164)/206 = 93%. Kappa = 0.74 (95% confidence interval: 0.68, 0.80).
Table 3. Numbers of Telephone Calls Required to Contact Participants for the Reliability Study, MOBILIZE Boston Study, Boston, Massachusetts, October 2007–February 2008

<table>
<thead>
<tr>
<th>No. of Calls Needed to Reach Participant</th>
<th>No. of Participants</th>
<th>% of Participants</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>119</td>
<td>57.8</td>
<td>57.8</td>
</tr>
<tr>
<td>2</td>
<td>51</td>
<td>24.8</td>
<td>82.6</td>
</tr>
<tr>
<td>3</td>
<td>18</td>
<td>8.8</td>
<td>91.4</td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>5.3</td>
<td>96.7</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>1.9</td>
<td>98.6</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>1.4</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>206</td>
<td></td>
<td>100</td>
</tr>
</tbody>
</table>

Abbreviation: MOBILIZE, Maintenance of Balance, Independent Living, Intellect, and Zest in the Elderly.

3-month recall correctly classified most participants who did not fall (164 persons by both methods), it missed one-quarter of those who did fall. Of 36 participants with a fall reported by the calendar method, 9 did not report a fall during the telephone survey, yielding a sensitivity of 75%. The specificity of the telephone survey for 3-month recall of falls was 96%. The chance-corrected agreement (kappa) was 0.74, indicating good agreement (19). The 9 participants who did not report a fall in the 3-month recall but did report a fall on their calendars had baseline characteristics that were similar to those of the overall study participants (e.g., similar age, sex, education, and MMSE score as compared with the variables in Table 2); however, at their baseline visit, they were likely to report 1 or more falls in the past year (6 of the 9 as compared with only 38% of the reliability study sample). No injurious falls occurred during our reliability study period, precluding an evaluation of whether severity of falls produced any difference in fall recall due to fall severity. Of the 36 participants who reported a fall on their calendars, 10 had more than 1 fall during the 3-month time period of the survey; however, consideration of these multiple falls did not affect the results.

Table 3 shows the numbers of telephone calls required to contact participants for this reliability study. We reached and completed the telephone survey for 206 members (95%) of the randomly selected subset population. We were unable to reach 12 participants (5%) within the 6-call limit. While more than half of the reliability sample was contacted with only 1 telephone call, the remaining 87 (42%) required 2–6 calls to complete the brief survey. Overall, an average of 1.97 calls were made per participant (standard deviation, 1.4). Over 90% of the sample was successfully contacted with 3 telephone calls. Similar to the overall parent cohort (in which up to 30% of participants required telephone clarification or a reminder), 28% of our reliability sample subjects required reminder phone calls to send in their calendars or to clarify data on the calendars.

DISCUSSION

In this study, although there was good agreement between the 2 methods when no fall had occurred, quarterly telephone recall resulted in a sensitivity of 75% in comparison with the calendar method. Thus, 25% of calendar falls were not reported by 3-month recall. Despite the additional staff resources needed for follow-up, falls calendars provided better information on falls. Thus, in cohort studies of older persons in which the occurrence of falls is a main outcome, prospective methods of falls data collection are preferable despite the extra calendar-related costs and personnel costs, due to concerns about underreporting. When using other, retrospective methods of falls reporting, investigators need to anticipate recall errors in reporting. In our study, 3% of participants reported a fall by quarterly recall who had not documented a fall on the monthly calendar. MacKenzie et al. (20) refer to this phenomenon as “telescoping,” in which a person may project a salient event that occurred outside of the recall period into that particular period. However, since both retrospective recall and prospective recording of falls by calendar method rely on self-reports, each method is susceptible to some source of error, which may be due to either forgetting a fall or “telescoping.”

In our study, we did not plan to examine subsets of respondents, such as those with lower cognition, those with multiple falls, or those with injurious falls. In fact, our sample did not provide us with sufficient statistical power to examine the impact of lower cognition on recall of falls. Cummings et al. (21) found that elders with lower MMSE scores were more likely to forget falls. The MOBILIZE Boston Study sample specifically excluded persons with low MMSE scores and had relatively few persons with MMSE scores between 18 and 23.

Prospective data collection with an aim of increasing the accuracy of information is labor-intensive, requiring contact by study staff, time for multiple telephone attempts at contact, etc. Nonetheless, we found that telephone recall methods also required time-intensive effort by research staff, evidenced by the high number of calls needed to reach participants in our reliability study.

There were several strengths and limitations of this study. A key strength is that the study population was composed of a population-based sample of community-dwelling elders, whereas other studies have involved self-selected samples (9, 10). In addition, the population included both men and women and the sample was racially/ethnically diverse, accurately representing older adults in the Boston area (15).

There were limitations to our study. We were not able to examine subsets within our data because of small numbers of events, despite having 206 participants. Additionally, we were not able to assess whether cognition affects reporting of falls by either method, since our sample did not include any moderately-to-severely cognitively impaired persons. However, both the MOBILIZE Boston Study cohort and the sample for this analysis included persons with lower cognition (MMSE scores of 18–23) who were able to provide consent and participate in the study and thus included more than the usual “healthy” aging population (12, 13).

It would have been useful to examine whether recall of falls over longer time periods (e.g., 1 year or various other time periods) was more accurate or yielded similar results. However, this was beyond the scope of our study because of staffing and cost limitations. In addition, we were unable to...
examine whether there was a difference in recall due to fall severity, since no injurious falls occurred during our 3-month study follow-up. Additionally, we did not collect data on economic costs, which might have been helpful in comparing potential costs (and cost savings) between the 2 methods of falls data collection.

Our reliability sample comprised a subset of persons taking part in a cohort study with falls as the main outcome. Thus, given the probable increased attention to recording falls each month, our reliability study participants may have been more aware of their falls in the past 3 months than another independent sample of older persons would have been. In addition, participants who filled out and returned their monthly falls calendars within the allotted time frame without a telephone prompt or reminder may have been more likely to have reliable quarterly recall of falls than those who failed to return their falls calendars and required reminder telephone calls. Thus, sensitivity and perhaps specificity in this study may have been greater than would be obtained in other situations (9, 10). Our study may actually have overestimated the reliability of the 3-month telephone recall approach. Finally, the details of a fall in a quarterly recall may not be as well remembered as the details of a fall that has occurred in the past month. We did not collect data on these aspects of recall.

An area for further study is the ascertainment of falls among persons who are cognitively impaired and whether there are differences according to level of impairment. In the Longitudinal Aging Study Amsterdam, persons who did not return their falls calendars were significantly older and performed worse on cognitive tests (22). Other studies have suggested that lower cognition resulted in worse recall of falls and that injurious falls were remembered better (9, 23, 24). However, literature regarding these topics is sparse, and since our study did not contain moderately or severely cognitively impaired participants, we are unable to comment on this aspect of falls reporting (10, 21).

In conclusion, this study, in agreement with other studies, has demonstrated that falls recalled in a quarterly telephone survey are likely to be underreported by elderly participants. In our reliability study, daily recording of falls using a falls calendar (prospective ascertainment) resulted in a greater number of reported falls than 3-month recall of falls occurring during the same time period (retrospective ascertainment), most likely indicating more accurate reporting of fall occurrence or underreporting of falls in the 3-month telephone recall. In summary, the monthly calendar approach detected more falls than telephone recall of falls in our population-based sample of community-dwelling older adults. Older persons may not remember falls as accurately when asked to recall them retrospectively over a 3-month period.

ACKNOWLEDGMENTS

Author affiliations: Institute for Aging Research, Hebrew SeniorLife, Boston, Massachusetts (Marian T. Hannan, Margaret M. Gagnon, Jasneet Aneja, Richard N. Jones, Lewis A. Lipsitz, Elizabeth J. Samelson, Douglas P. Kiel); Harvard Medical School, Boston, Massachusetts (Marian T. Hannan, Richard N. Jones, Lewis A. Lipsitz, Elizabeth J. Samelson, Douglas P. Kiel); Department of General Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts (Marian T. Hannan, Richard N. Jones, Lewis A. Lipsitz, Elizabeth J. Samelson, Douglas P. Kiel); Department of Biostatistics, School of Public Health, Boston University, Boston, Massachusetts (L. Adrienne Cupples); Framingham Heart Study, Framingham, Massachusetts (L. Adrienne Cupples); and College of Nursing and Health Sciences, University of Massachusetts, Boston, Massachusetts (Suzanne G. Leveille).

This work was supported by the National Institute on Aging (Research Nursing Home Program Project grant P01AG004390 to L. E. L. and research grants R01AG026316 to M. T. H. and R37AG25037 to L. E. L.).

The authors acknowledge the MOBILIZE Boston Study research team for their time, effort, and dedication. They also thank Marcie Freeman for her recruitment and outreach efforts and Christopher Rockett for his analytic programming.

This paper was presented at the 61st Annual Scientific Meeting of the Gerontological Society of America, National Harbor, Maryland, November 21–25, 2008.

The sponsors had no direct role in the conduct of this study.

Conflict of interest: none declared.

REFERENCES