Patient Recall of Health Care Events and Time to Diagnose a Suspected Ovarian Cancer

Lisa M. Hess,1 Michael W. Method,2 Frederick B. Stehman,1 Tess D. Weathers,1 Paridha Gupta,1 Jeanne M. Schilder1

Abstract

This study sought to explore the consistency of patient recall compared with data found in health care records in the setting of referral to a gynecologic oncologist for suspected ovarian cancer. Ninety-one patients were interviewed regarding symptoms and health care events, and the dates leading to diagnosis and medical record histories were obtained and abstracted. There was low correspondence between patient recall and the medical record for time to diagnosis and health care events. This study emphasizes the need to find a way to obtain information about patient history and test results directly from providers in a timely manner, as patients often do not even recall the health care events or tests they have received.

Introduction:
Patient recall is often used by clinicians to create a history of care leading to consultation with a gynecologic oncologist. Although patient recall may be an efficient method to explore the context of the patient’s concerns, the accuracy of recall and its potential impact on care are unknown. This study sought to explore the consistency of patient recall compared with data found in health care records. Patients and Methods: This study enrolled 105 eligible patients who were referred to a gynecologic oncologist for suspected ovarian cancer. Ninety-one of these patients were interviewed regarding symptoms, health care events, and the dates leading to diagnosis. The medical records of these patients from all previous providers were obtained and data were abstracted. The intraclass correlation coefficient (intraclass correlation coefficient, ICC[3,1]) was used to examine correspondence between recall and medical record data. Results: There was low correspondence between patient recall and the medical record for time to diagnosis (ICC = 0.12; 95% confidence interval [CI] = −0.09 to 0.33; P = .12) and health care events (0.15; 95% CI, −0.05 to 0.348; P = .008). Conclusions: There should be limited confidence in information obtained from patient recall given the inconsistency between recall and actual dates and events.

Clinical Ovarian & Other Gynecologic Cancer, Vol. 5, No. 1, 17-23 © 2012 Elsevier Inc. All rights reserved.

Keywords: Health care events, Intraclass correlation, Recollection, Reliability, Time to diagnosis

Introduction
An estimated 22,280 women were diagnosed with ovarian cancer in the United States in 2012, and 15,500 died of the disease.1 In a recent study by Bailey et al.,2 30% of all patients referred to a gynecologic oncologist for suspected ovarian cancer have nonsignificant findings at surgery. For the remaining 70%, prompt diagnosis and surgical intervention is crucial for long-term survival.3

A potential barrier to a rapid diagnosis in the absence of screening tests for ovarian cancer is that patients may undergo multiple clinic visits and repeated assessments over an extended time to rule out other conditions before referral to a gynecologic oncologist. An Australian population-based telephone survey sought to examine the pathways to diagnosis of ovarian cancer to explore this possibility.4 They found that women diagnosed with ovarian cancer experience an average of less than 2 health care visits, during which only 10% had their cancers diagnosed incidentally. The results from this study suggest that women who present with symptoms are diagnosed relatively quickly (80% self-reported that they were diagnosed within 3 months of initial presentation).4 This was slightly longer than the average one-month time to diagnosis recalled in an earlier Canadian self-report study.5 However both studies relied on patient recall of past events, and the reliability of recall in the setting of suspected ovarian cancer is unknown.
Patient Recall of Health Care

There is some debate regarding the reliability and accuracy of patient recall in both the short term and long term for a variety of conditions. Some studies have found patient recall of health care events to be reliable (correlation coefficients ranging from 0.6-1.0), whereas others have found that inaccurate recall of symptoms and events is common, with up to 26% of patients who cannot accurately recall previous health care events and procedures. We are unaware of published research on the reliability or accuracy of patient recall of symptoms or health care tests or events in the setting of the diagnostic pathway to ovarian cancer. Therefore, there is a need to explore the accuracy of patient recall of both health care events (previous tests and visits) and time leading to diagnosis to better understand the clinical implications of relying on patient recall.

This study was designed to test the hypothesis that patient recall is consistent with medical record data regarding health care events and time to reach an accurate diagnosis. This evidence will enable us to determine the confidence that we can have regarding the assumptions or conclusions that are made based on recall for clinical care in this setting.

Patients and Methods

To be eligible for participation, participants were required to be 18 years of age, to not yet have undergone surgery but to have received a referral for surgery, or to have undergone surgery within the past 12 months for a suspected ovarian malignancy. Women with a diagnosis of ovarian cancer or a definitive surgical procedure that was completed more than 12 months before enrollment were not eligible in order to reduce recall bias and to be consistent with earlier studies of symptoms and pathways to diagnosis.

After providing written informed consent, eligible women participated in a semistructured interview that was audio recorded to ensure data collection accuracy. The interview was designed for the patient to first identify the situation or symptoms that prompted seeking medical attention and to recall the experience from that time to the time of surgery for the suspected ovarian malignancy. Variables collected during the interview process included the date of each health care visit, physicians and other health care providers consulted, tests, prescriptions, and referrals for each health care system encounter/clinic visit from the time that a problem was first suspected by the patient through the time of definitive pathologic diagnosis as defined as the date of definitive surgery with pathologic diagnosis. In the case of biopsy with pathologic diagnosis without definitive surgery, the biopsy date was used.

The theoretical framework for the diagnostic pathways is based on Anderson’s model of total patient delay. This model is composed of a series of patient and health care system factors, beginning with the initial inference of illness from noted symptoms, leading to the patient seeking medical attention and ultimately obtaining an accurate diagnosis. For this study, the appraisal period is defined as beginning at the first symptoms or health concern detected by the patient or friends/family (whichever is earlier) and ending on the date of the first medical encounter at which the problem was noted by the patient or provider. The diagnostic period is defined as the time from the first medical encounter at which the health concern was noted by the patient or provider through the time of definitive diagnosis. Definitive diagnosis is defined as a pathologic or histologic diagnosis as described above. The time-to-treatment period is defined as the time from definitive diagnosis to time of chemotherapy (limited to patients diagnosed with cancer who are recommended to receive chemotherapy). All study procedures were approved by the Indiana University/Clarian Health and Northern Indiana Research Consortium Institutional Review Boards.

The target accrual goal was 100 women to attempt to reach a sample of 30 patients in each diagnostic category (ie, benign disease, early-stage disease, and advanced-stage disease), which would allow for 80% power to detect a correlation of 0.50 or greater in each category and 80% power to detect low correlations (0.30) between self-reports and the medical record data among women diagnosed with ovarian cancer.

All analyses were conducted using SPSS, version 15.0 (SPSS Inc, Chicago, IL). Descriptive statistics were used to present participant age, marital status, insurance status, race, and diagnosis of study participants. For comparative analyses, racial categories were collapsed into white/nonwhite categories because of the low recruitment of Asian, Native American, and African American women and women of other races. Reported symptoms, cancer vs. benign diagnosis, and cancer stage at diagnosis were compared for differences in time periods using analysis of variance, with posttests adjusted for multiple comparisons in the case of an overall significant result. Mean differences in time periods by presence of symptoms, insurance status, and cancer diagnosis vs. benign disease were conducted using independent sample t tests, and correlations were performed to explore the relationship between age and time/events.

To assess the agreement of number and timing of health care events between patient recall and the medical record, the intraclass correlation coefficient (ICC) was used to evaluate the agreement between raters (in this case between the self-report and the medical record). Because of the nonrandom nature of these assessments, the Spearman-type ICC form (Model C) was used. The ICC Model C (ICC[3,1]) was analyzed as a 2-way mixed effects Model ICC using SPSS, version 15.0 (SPSS, Inc). If a significant correlation coefficient was found, a comparison of means was performed to better interpret the significant findings.
For consistency with previous research, which has reported number of months for the appraisal and diagnostic periods, patients who failed to recall time within 30 days of that found in the medical record for the illness appraisal, diagnostic period, or time-to-treatment period were considered to have inaccurate recall for those time periods. For the number of events, patients who failed to recall 3 or more health care visits were considered to have inaccurate recall, which is very conservative and equivalent to the total number of visits experienced by most patients from earlier research.

### Results

One hundred seven patients were enrolled in this study. One patient was excluded because more than 12 months had passed since the diagnosis, and 2 additional patients changed their minds about participating before scheduling a study interview. Of the 104 eligible patients, 92 completed the study interview. Thirteen patients did not complete the interview because they could not be reached by phone or in person. One patient who was interviewed was excluded from the analysis because all health care encounters occurred outside the United States, resulting in a sample size of 91 patients who had complete medical record data.

### Analysis

#### Illness Appraisal Period

Four patients (4.4%) identified becoming aware of potential illness after first seeking medical care, so the illness appraisal period was set to zero according to patient recall. However according to the medical record, 27 (29.7%) patients had seen a health care provider related to these concerns before the time they reported as first suspecting illness. According to patient recall, the illness appraisal period was an average of 72 days (range 0–1096 days).

There was a significant positive relationship between age and the illness appraisal period \(P = .02\) according to patient recall. Patient recall of appraisal could not be verified in the medical record, since this event occurred independently of any medical care obtained or received.

#### Diagnostic Period

Three patients could not recall the date of diagnosis and were excluded from the comparison and correlation analyses. Overall, pa-
There was a mean difference of 71.8 days between patient estimation of the time to diagnosis compared with that found in the medical record (range of an underestimation of 181 days to an overestimation of 3430 days). The ICC between the patient-recalled and medical record diagnostic time period was 0.122 (95% confidence interval [CI], −0.085 to 0.329; \( P = .12 \)), demonstrating no agreement. More than a quarter (27.3%) of patients could not accurately recall the diagnostic period within 30 days.

**Time-to-Treatment Period**

Of the 63 patients who were diagnosed with cancer in this study, 58 planned to have or had received chemotherapeutic treatment. Six patients had not yet decided on the start date of chemotherapy at the time of the interview or did not know the date of diagnosis and were excluded from the comparative analyses because of the inability to calculate a time-to-treatment period for patient recall. The patient-recalled time-to-treatment period was an average of 34.6 days (range 0-162 days); in the medical record this period was 33.2 days (range 0-132 days). There was very high consistency between the patient-recalled and medical record time-to-treatment periods, with an ICC of 0.808 (95% CI, 0.687-0.886; \( P < .001 \)), demonstrating high agreement. There were no significant differences between these 2 data collection methods, with a mean difference of 3.8 days (range of an underestimation of 48 days to an overestimation of 64 days). The vast majority of patients were able to accurately recall the time-to-treatment period, with only 3 patients (5.9%) unable to accurately recall this period within 30 days.

There was a significant difference according to patient recall based on insurance status, with uninsured patients reporting an average of 56.2 days from diagnosis to treatment vs. insured patients reporting an average of 31.8 days (\( P = .03 \)); however there was no significant difference according to the medical record data. There were no significant differences in time to treatment by race or marital status. There was no difference in time to treatment by age according to patient recall; however in the medical record there was a significant relationship between age and time to treatment, with older patients receiving treatment more rapidly than younger patients (\( P = .02 \)).

**Number of Health Care Visits**

Patients recalled having an average of 3.4 health care visits (range, 1-8), and according to the medical records patients had an average of 7.1 health care visits (range, 2-27) during the diagnostic period (Table 4). The ICC between the patient-recalled and medical record events was 0.149 (95% CI, 0.002-0.348; \( P = .001 \)), demonstrating low agreement. Patient recall reported significantly fewer events than did the medical records (\( P < .001 \)). On average, patients underestimated the number of health care visits by 3.7 (range, underestimation of 20 events to overestimation of 4 events). Fifty-two participants (57.1%) were inaccurate by 3 or more visits when all visits, including tests and assessments, were included.

When health care visits involving testing and assessments only (eg, no health care provider was seen), there remained a low agreement between patient recall and the medical record (ICC = 0.268; 95% CI, 0.002-0.487; \( P < .001 \)). There was an average of 3.1 visits (range, 1-8) according to patient recall and 5.1 medical record events...
(range, 2-19) when all testing visits were excluded. There were a significantly lower number of recalled events than medical record events \((P < .001)\), with an average underestimation of 2 medical visits (range, underestimation of 14 events to an overestimation of 4 visits). Twenty-eight patients (30.8%) could not recall 3 or more health care visits when all testing and assessment visits were excluded from the computation.

When comparing the number of events by patient characteristics, insured patients \((P = .05)\) and patients with symptoms \((P = .01)\) reported significantly more health care events (including testing) than did uninsured patients or patients without symptoms, respectively; however this was not different according to the medical records or when tests were excluded in the patient self-report. There was no difference in number of health care visits between patients who were later diagnosed with cancer and those with benign disease or by marital status, race, or age.

**Providers and Tests**

According to patient recall, there were a total of 313 health care visits during the diagnostic period, including 28 visits for testing only. In the medical record, there were 648 visits, including 179 for testing only (Table 5). The most common practitioners seen for the initial presentation, according to patient recall, were primary care physicians \((n = 50 [54.9\%])\), emergency room or urgent care \((n = 18 [18.8\%])\), or gynecologists \((n = 8 [8.8\%])\). Similarly, in the medical record it was found that 46 women \((50.5\%)\) first presented to a primary care physician, 24 \((26.4\%)\) presented to the emergency room or an urgent care facility, and 8 \((8.8\%)\) presented to a gynecologist. Of the 63 patients who were diagnosed with cancer, 47 \((74.6\%)\) reported seeing a gynecologic oncologist before diagnosis. According to the medical record, 49 \((77.8\%)\) patients who were diagnosed with cancer saw a gynecologic oncologist before pathologic diagnosis. Participants were least likely to recall visits for testing only or visits to a gastroenterologist and were most likely to recall visits to a gynecologist, gynecologic oncologist, or urologist. Patients who saw a gynecologic oncologist during the diagnostic period recalled a significantly shorter time to diagnosis \((P < .001)\); however there was no difference in time to diagnosis according to the medical record. There was no significant difference in the diagnostic period depending on where the patient initially presented for care. However when outliers were removed (patients with longer than a 3-year diagnostic period), those first presenting to primary care physicians, gynecologists, or the emergency room reported a significantly shorter time to diagnosis \((P < .001)\); however there was no difference in time to diagnosis according to the medical record when outliers were removed. There was also no significant difference in the number of health care visits during the diagnostic period, depending on the provider seen at initial presentation, if the patient was currently receiving chemotherapy, or based on time between diagnosis and study interview.

**Factors Related to Recall**

There was no difference in accuracy of the recalled periods (diagnostic period or time to treatment) according to insurance status, cancer diagnosis, initial symptoms, race, or marital status. However, patients who inaccurately recalled the diagnostic period were more likely to be older. Patients with accurate recall were an average age of 55 years and those with inaccurate recall of the diagnostic time period were an average age of 63.3 years \((P = .007)\). Age was not associated with accuracy of recall for the illness

### Table 4: Visits to Health Care Professionals During the Diagnostic Period, Number of Patients (%)

<table>
<thead>
<tr>
<th>Number of Visits</th>
<th>Patient Recall</th>
<th>Medical Record Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All Visits</td>
<td>Excluding Visits for Tests/Assessments</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1 (1.1%)</td>
<td>2 (2.2%)</td>
</tr>
<tr>
<td>2</td>
<td>28 (30.8%)</td>
<td>34 (37.4%)</td>
</tr>
<tr>
<td>3</td>
<td>2 (28.6%)</td>
<td>30 (33%</td>
</tr>
<tr>
<td>4</td>
<td>20 (22%)</td>
<td>14 (15.4%</td>
</tr>
<tr>
<td>5</td>
<td>7 (7.7%)</td>
<td>4 (4.4%)</td>
</tr>
<tr>
<td>6</td>
<td>4 (4.4%)</td>
<td>3 (3.3%)</td>
</tr>
<tr>
<td>7</td>
<td>1 (1.1%)</td>
<td>1 (1.1%)</td>
</tr>
<tr>
<td>8</td>
<td>4 (4.4%)</td>
<td>3 (3.3%)</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15+</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mean (Range)</td>
<td>3.4 (1-8)</td>
<td>3.1 (1-8)</td>
</tr>
</tbody>
</table>
appraisal or for the time-to-treatment periods. Patients with a longer appraisal period (according to the medical record) were more likely to inaccurately report that period, but there was no difference in either the diagnostic period or time-to-treatment period between those who reported inaccurate time periods and those who reported them accurately.

There was no difference in accuracy of the recalled periods (diagnostic or time to treatment) according to insurance status, cancer diagnosis, initial symptoms, race, marital status, age, current chemotherapy exposure, or time from diagnosis to study interview. Accuracy of recall of visits was significantly related to the number of visits that occurred during the diagnostic period. Patients who accurately recalled their health care visits (did not miss more than 2 visits), had an average of 4.6 visits, whereas patients who failed to recall 3 or more events had an average of 9 health care visits during the diagnostic period ($P = .03$). This relationship was similar when the visits for tests only were removed from the analysis, with an average of 3.9 visits for patients with accurate recall vs. 7.9 events for those who could not recall 3 or more events ($P < .001$).

## Discussion

Patients referred to a gynecologic oncologist for suspected ovarian cancer did not accurately recall their health care experience leading to referral and diagnosis. Although patients could recall time periods (generally within 30 days), they were not as likely to recall all the health care visits that occurred during that period. Although these study participants were recruited from Indiana, the patients (demographic and clinical variables) and their self-reported data (time and events) in this study are very similar to those in the Goff symptom$^{10}$ and Australian pathway$^{4,19}$ studies. For example, patients in this study reported an average of 3.4 visits to reach diagnosis, similar to the 2.6 visits reported in the Australian study.$^4$ However in the medical record, there was an average of 7.2 visits, and it was found that 58% of patients failed to recall 3 or more visits. Patients reported an average of 4 months to reach diagnosis...
in this study, similar to the 4.4 months reported in the Goff symptoms study.\textsuperscript{4,19} However, when compared with the medical records it was found that patients underreported the time to diagnosis on average by more than 1 month. This suggests that the pathway to diagnosis of ovarian cancer uses many more health care system resources that had previously been reported.

It is unlikely that the medical record data are affected by the retrospective nature of data collection, as the number and type of events, providers, and tests should not be affected. It would be practically impossible to conduct a similar study prospectively because the sample would necessarily require many thousands of healthy women to find up to 100 women who might possibly suspect a malignancy in the future. However because of the retrospective nature of data collection, there is the possibility of missed events or tests in the medical record, which would only serve to increase the discrepancy between recall and actual events. There is also the possibility of better recall in this study because of the detailed survey procedures used to elicit details of each test, visit, and physician seen throughout the diagnostic period. In clinical practice, it is possible that less detailed methods are used to create a recent medical history, that less detailed methods are used to create a recent medical history, and this study may underestimate recall in a natural setting.

The study population was limited to patients in Indiana and may not be generalizable to the larger population with ovarian cancer. However both private practices and an academic site recruited patients to enhance the representativeness of study participants. Additionally, the demographic and clinical characteristics and the self-reported data are very similar to the data collected in other population-based and recall trials (including mean age, disease stage, and percent reporting symptoms),\textsuperscript{4,16} suggesting that this study population is comparable to those in previous work.

Medical record review is complex and challenging, particularly in a fragmented health care system. Time and resources are needed to identify, obtain, and record health care system events, resources, and practitioners involved in the diagnostic pathway. The use of a claims database is a potential solution but may affect the outcomes because of the specific guidelines and recommendations for tests and referrals within that particular insurance system. However, the results of studies that rely on claims data may not be generalizable to all payers.

**Conclusion**

Patients with suspected ovarian cancer did not accurately recall health care events or time leading to referral to a gynecologic oncologist and diagnosis. In this study, 60% of patients were inaccurate in their recall by more than 3 health care events, and approximately 10% were inaccurate by more than 7 events. More than 27% of patients did not recall the period in which the events occurred within 30 days, and approximately 15% could not recall the time within 70 days. Only 5.5% were highly accurate in the correct recall of number of health care events, and 40% were highly accurate regarding time (within 1 week of what was found in the medical record).

Patients are most likely to fail to recall tests that have been performed. It will always be necessary for some research to be conducted that relies on patient self-reporting of time or events. Based on the findings of this study, research using patient recall data should take into account the uncertainty and error in self-reported data and perhaps adjust conclusions based on the potential range of inaccuracy. Conclusions from self-reports should not assume that there is no uncertainty in the results.

**Clinical Practice Points**

- If clinical decisions are to be made based on previous health care events, tests, or time to referral, the discrepancy between patient recall and medical record data cannot be ignored.
- Medical record review is recommended for research, particularly if recommendations for care will depend on results obtained from patient-recalled time or event data.
- Clinicians should recognize the inaccuracy of patient recall of health care events and time in the setting of referral to a gynecologic oncologist for suspected ovarian cancer.

**Acknowledgments**

This study was supported by an internal pilot grant of the Indiana University, Mel and Bren Simon Cancer Center, funded by the American Cancer Society Institutional Grant program (IRG-84-002-22).

**Disclosure**

The authors have stated that they have no conflicts of interest.

**References**