SURVEY DESIGN CHALLENGES IN SAMPLING RECENTLY DIAGNOSED CANCER SURVIVORS

Maxine Denniston, American Cancer Society, Donna Brogan, Emory University, Frank Baker, Joseph Dye and Danette Hann, American Cancer Society
Maxine Denniston, American Cancer Society, 1599 Clifton Road, NE, Atlanta, GA 30329

Key Words: Sampling frame, Survey design, Cancer survivors, Frame construction

Abstract

The Behavioral Research Center of the American Cancer Society is conducting a nationwide population-based study of cancer patients focusing on quality of life. The population of inference is U.S. adults diagnosed with 1 of 10 cancers within a given 12-month period. The goal is to obtain baseline data 6-12 months after diagnosis. Since most cancer patients are older, oversampling is planned for younger patients to provide an adequate sample size to study issues relevant to this age group (e.g., employment, medical insurance). Racial minorities will also be oversampled. Data collection is via self-administered mailed questionnaires.

The strategy was to develop a single stage stratified sampling plan in collaboration with each state cancer registry and if this approach was not feasible, to conduct 2-stage cluster sampling of facilities where cancer is diagnosed followed by stratified sampling of cases. The 2-stage cluster sampling approach would be administered either through the state cancer registry or by another local entity such as an academic institution. Our experience in 12 states illustrates issues likely to be faced in additional states. Examples are: (1) In most states it is not possible to develop a sampling frame of patients based upon the routine operation of the cancer registry because data on newly diagnosed cases are not reported quickly enough. (2) Patient confidentiality requires that registry personnel obtain informed consent from sampled patients. Also, in most if not all states, registry personnel must obtain consent from the patient's physician before his/her patients can be contacted regarding the study. These processes may differ for each state. (3) In some states, it is not legally permissible to let patient contact information leave the state, even with patient consent. This paper discusses some of the design and fieldwork challenges of this sample survey and a redesign of the study based on our experience.

Introduction

Today in the U.S., over 10 million individuals are alive who have been diagnosed with cancer and over 7 million have survived their cancer for 5 or more years[1]. Little, however, is known about the long-term needs of cancer survivors and their families. Recently research interest has focused on factors related to quality of life in cancer survivors. It is convenient to sample incident cancer patients from large clinical facilities that specialize in cancer diagnosis and treatment; however, patients who are treated at such facilities most likely are not representative of all incident cancer patients. Additionally, much of the existing research has focused on patients who are still in their initial course of treatment (e.g. surgery, radiation, chemotherapy and/or hormonal therapy) and does not follow patients after treatment is completed.

The Behavioral Research Center of the American Cancer Society (ACS) is conducting a study of quality of life in a sample of cancer survivors that would be representative of all cancer survivors. There are several definitions of the term "cancer survivor", but the one adopted by ACS for its proposed study was that of the National Coalition of Cancer Survivors, which states that a cancer survivor is "an individual who is living and has been diagnosed with cancer". The goals of the study are (1) to identify needs of cancer survivors and their families, (2) to test hypotheses concerning variables that affect quality of life and good survivorship, (3) to provide a database for planning and evaluating ACS programs for survivors and (4) to examine factors that relate to late effects of cancer and its treatment, including second cancers.

Initial Study Design

The initial plan called for a nationwide, longitudinal study that would include up to 100,000 selected incident cancer patients. This sample size represented approximately 10% of the 909,100 newly diagnosed cases of the cancers to be included in the study that were estimated to occur in 1998[2]. The population of inference was to be adult (18 and older

at time of diagnosis) cancer patients who had been diagnosed with 1 of the 10 most common cancers (prostate, female breast, lung, colorectal, bladder, Non-Hodgkin's lymphoma, skin melanoma, kidney, ovarian and uterine)[1] within a given 12 month period. For ACS organizational and program reasons, it was planned to conduct the study in all 50 states. Hence, each state is considered to be a stratum. Population based samples were to be selected from incident cases in a given 12 month period for each of the 50 states. The sample in each state was to be of sufficient size for state level analysis, so that results could be used for advocacy by local ACS units. Since most cancer patients are older, oversampling was planned for patients aged 18 to 54 in order to provide an adequate sample size to investigate issues related to medical insurance, employment, raising dependent children, etc. Racial minorities and cancers with lower survival rates were also to be oversampled. Baseline data on participants was to be obtained within 6-12 months of diagnosis, with follow-up surveys 1, 3 and 5 years after baseline and possibly up to 10 years after diagnosis. The 6-12 month post-diagnosis window for baseline data collection is critical because it is hypothesized that important psychosocial adjustments take place in that time period that will determine later quality of life and reintegration back to a "normal" life. We expected a 50% yield rate, that is, 50,000 completed baseline questionnaires from 100,000 selected patients. This expected yield rate was based on anticipated non-response at the physician and patient consent stage and questionnaire completion, as well as anticipated mortality during the time between selection into the sample and completion of questionnaires, especially among lung cancer cases.

Thus the general outline of the study process was as follows:

1. Each state cancer registry would develop, negotiate and finalize a work plan and budget to perform the work requested by the ACS.
2. Each state cancer registry would obtain Institutional Review Board (IRB) or similar peer review approval for the study.
3. Each state cancer registry would construct the sampling frame of its newly diagnosed cases. This would be accomplished either through its standard operating procedure, through rapid case ascertainment or through 2-stage cluster sampling of facilities followed by stratified sampling of cases.
4. Cases would be stratified by the 10 types of cancer included in the study, by age category (18-54 and 55 or older) and in some states by race or other demographic characteristic (e.g., residence in an urban or rural area).
5. A probability sample of cases would be selected using sampling fractions developed by the ACS biostatistician. Varied sampling probabilities would reflect oversampling of cancers with lower survivals, patients aged 18-54 and, where applicable, minorities.
6. Registry personnel would obtain consent from the physician of each sampled patient to contact his/her patient(s) regarding the study. The specified physician consent procedure included both a second mailing and telephone follow-up with non-respondent physicians.
7. After obtaining physician consent, registry personnel would contact each sampled patient to introduce the study, to ask him/her to participate in the study and to obtain informed consent to release the patient's contact information to ACS. The specified consent procedure for patients also included a second mailing and telephone follow-up with non-respondent patients.
8. Names, addresses and telephone numbers of consented patients would be sent periodically by the cancer registry to the ACS contractor who would be responsible for questionnaire distribution, receipt of completed questionnaires and all subsequent steps up to and including the provision of raw data files to the ACS Behavioral Research Center. The questionnaire distribution procedure included a second mailing and telephone follow-up with non-respondent patients.

Sampling Frame

A nationwide survey using some form of random digit dialing or area probability sampling was considered not economically feasible because this subpopulation (i.e., newly diagnosed cancer patients with 1 of the 10 target cancers) constitutes a very low percentage of the general U.S. population (estimated

to be < 0.5% based on 909,100 incident cases estimated to occur in 1998[2] and an estimated U.S. population aged 18 and over of 200,426,000 for July 1998[3]). Theoretically, a sampling frame for the proposed study does exist through state and SEER (Surveillance, Epidemiology and End Results) cancer registries. In most cancer registries, it takes about 2 years from the beginning of a given calendar year for all cases diagnosed in that year to be reported to the registry. To obtain baseline data from study participants within 6-12 months of diagnosis and to have reasonably complete sampling frame coverage would require a registry to have 90% of its cases reported within 9 months of diagnosis. The timeliness of routine reporting of cases to the cancer registries is not usually sufficient to allow for baseline data to be obtained within the desired timeframe. In most states it is necessary to use a method known as rapid case ascertainment in order for the baseline questionnaire to be administered in the 6-12 month post-diagnosis window. Rapid case ascertainment is, however, more costly and complicated than the routine reporting system, and is likely to vary from state to state. In addition, not all cancer registries currently have a rapid case ascertainment system in place. Thus, no standardized protocol for case identification can be imposed in all 50 states; rather, the protocol will need to be modified to fit each cancer registry's operations.

Legal Issues

In each state, IRB or similar peer review approval for the study must be obtained with the aid of the cancer registry. This is a time consuming process that varies from state to state. Even with IRB approval, however, state laws prohibit the cancer registries from releasing contact information on individual cases without first obtaining informed consent from the patient. In some states, it is not legally permissible to let patient contact information leave the state, even with the patient's consent. Also, in most if not all states, the patient's physician must first give consent for his/her patients to be contacted regarding the study. In addition to being a legal requirement, obtaining physician consent also functions as a professional courtesy and has the advantage of providing an opportunity to identify patients who are ineligible for the study (i.e., patients who have died since being diagnosed or who have known psychiatric or neurological disorders). The form of physician consent varies from state to state, with some states requiring that the physician respond that it is okay to contact each patient (active consent), while other states allow patient consent to proceed if a negative response is not received from the physician within a specified time period (passive consent).

Once physician consent has been obtained for a sampled patient, registry staff must then obtain informed consent from each sampled patient for study participation and for release of patient contact information to the ACS or its designee. Registries require a signed written consent before releasing patient contact information to a research organization such as ACS. A verbal consent at telephone follow-up with non-respondents to the second mailing that does not result in receipt of the written consent form is not considered legal consent. Note that, due to legal considerations, the cancer registry must successfully complete all the work of both the physician and patient consent before the ACS can make any contact with a patient. Names, addresses and telephone numbers for consented patients are periodically sent by the registry to the ACS contractor who is responsible for questionnaire distribution, receipt of completed questionnaires, and subsequent steps in the study up to and including the provision of raw data files to the ACS Behavioral Research Center.

Registries are reimbursed for the time and materials they spend on rapid case ascertainment, sample selection, and physician and patient consenting. This requires that each registry prepare and submit a state-specific work plan and budget that is subject to multiple rounds of review at both the registry and the ACS before final approval. This process will vary in its complexity and timeframe from state to state.

Challenges

The challenges are largely concerned with the work done by the cancer registries and the ACS contractor; that is, the team conducting the study (i.e., ACS Behavioral Research Center) has less than optimal control over the process. Much of this loss of control arises from legal issues. In general, it is impossible for the ACS to obtain patient contact information without the patient's consent to do so; thus ACS must work through the cancer registries to obtain participation in the study from sampled patients. This means that the work to be carried out by the registries must be carefully specified and monitored by the ACS. Unfortunately this monitoring must be done from off-site, which is more difficult than monitoring work that is done locally. Because of these legal considerations, it is harder to obtain a high overall survey response rate. First, permission must be obtained from each patient's physician to contact

his/her patient(s) regarding participation in the study. Then the patient must return a signed consent form before a questionnaire can be mailed, instead of using the implied consent that is often assumed in mail surveys. Finally a completed questionnaire must be returned. Thus, there are 3 points instead of 1 at which an individual can refuse, but this process is required by legal constraints. There are also ethical considerations, since the consent process does provide 2 points at which the death of a selected patient could be discovered before the mailing of a questionnaire, the arrival of which may upset surviving family members. Other challenges arise from the desire to obtain baseline information within 6-12 months of diagnosis (i.e., problems related to the need for the registries to use rapid case ascertainment methods).

Lessons Learned

Some valuable lessons have been learned from piloting the study in 3 states and from working with several other states in preparation for study implementation. First, for a number of reasons, it will not be possible to do the study in all 50 states. We have learned that a much larger than anticipated amount of time and work is required in each state for IRB approval and contract negotiations. We have found that even in a state where 90% of cases are routinely reported to the registry within 9 months of diagnosis, events such as a change in the computer software used for reporting can delay case identification to the point where many cases are not sampled in time to meet the goal for baseline data collection within 6-12 months of diagnosis. Additionally, we have discovered that many registries will not be able to participate in the study at all due to operational considerations (e.g., no rapid case ascertainment system in place, time and staffing considerations) or to current and future involvement in numerous other research projects utilizing patients from their registry. There are a few states that until very recently did not have a cancer registry and there are still states in which the existing registry does not have adequate statewide coverage (including the recently established registries). It was originally thought that an approach using 2-stage cluster sampling of facilities followed by stratified sampling of cases, administered through a local entity such as an academic institution, could be used to identify cases in these states. This approach has been tried in Georgia, with the ACS Behavioral Research Center being the local entity that administers the study. The approach, however, has proven to be cumbersome and impractical due to legal issues of patient confidentiality and the need to recruit and obtain IRB approval at each of the 40 sampled facilities (clusters).

Second, we have found that patient consent and questionnaire return rates are lower than desired for scientific validity. Before fielding the study in additional states, we must determine the reasons for the low patient consent and questionnaire return rates and modify the applicable elements of the study (e.g., more effective study introduction letter, shorter questionnaire) to obtain higher levels of participation.

Survey Redesign

The lessons from the pilot states make it clear that the study plan must be redesigned in order to be feasible. First, there will be a purposeful selection of cancer registries to be recruited for study participation. These selected registries should provide a reasonably representative geographic coverage of the U.S., should contain sufficient numbers of minority populations and should include both urban and rural areas. Second, it will still be necessary to work through the cancer registries for frame development and sample selection. Registries also will continue to obtain physician and patient consent, but more detailed procedures, that are as uniform as possible, must be provided for the consent processes, and ACS must more closely monitor the performance of the consent procedures by the registries. Third, it will rarely if ever be possible for case ascertainment to be carried out via a registry's routine case reporting procedures; rather some sort of rapid case ascertainment will need to be utilized. If rapid ascertainment of all its cases is not feasible due to staffing constraints or if it is too expensive for a state to perform, the possibility of using 2-stage cluster sampling of facilities followed by stratified sampling of cases will be explored. Thus, cases would only need to be identified using the more costly and labor intensive rapid ascertainment from a portion of the facilities that report to the state registry. Finally, this alternative approach using 2-stage cluster sampling of facilities followed by stratified sampling of cases will only be attempted in collaboration with a state cancer registry as a cost saving or labor saving measure or to increase the statewide coverage of the registry. It will not be viewed as a method to obtain data through an entity other than the state cancer registry.

Conclusions

It is never easy to do population based research. In a setting with legal issues, concerns for

the feelings of the family of deceased individuals and the desire to obtain data in relatively close proximity to a diagnosis of cancer, the job becomes even more difficult.

References

1. *Cancer Facts & Figures – 1996.* American Cancer Society, Atlanta, GA, 1996.
2. *Cancer Facts & Figures – 1998.* American Cancer Society, Atlanta, GA, 1998.
3. Population Estimates Program, Population Division, U.S. Bureau of the Census, Washington, DC, October 1, 1999.