Invited Commentary: Birth Certificates—A Best Control Scenario?

Julie A. Ross1,2, Logan G. Spector1,2, Andrew F. Olshan3, and Greta R. Bunin4

1 Division of Epidemiology and Clinical Research, Department of Pediatrics, University of Minnesota, Minneapolis, MN.
2 University of Minnesota Cancer Center, Minneapolis, MN.
3 Department of Epidemiology, University of North Carolina, Chapel Hill, NC.
4 Children’s Hospital of Philadelphia, Philadelphia, PA.

Received for publication January 16, 2004; accepted for publication February 12, 2004.

Abbreviation: RDD, random digit dialing.

The majority of conditions that affect children under the age of 15 years (e.g., birth defects, diabetes, cancer) are sufficiently rare to warrant the case-control study approach. Furthermore, most pediatric studies require enrollment from multiple centers, registries, or hospitals to gain a sufficient sample size and thus can span large geographic areas. While selection of cases is usually straightforward, special care must be given to selection of controls to ensure that they represent the underlying population from which the cases arose. As Wacholder et al. (1–3) nicely described in a 1992 series of articles addressing control selection issues, an underlying cohort or study base can be primary or secondary. The primary base is the hallmark of a population-based case-control study in which all cases in a defined population are ascertained both geographically and temporally. When it is difficult to ascertain all cases in an area, a secondary base is chosen. In this instance, the base is defined as the cases, and the controls are identified as persons who would have become study cases had they developed the disease during the time period. For example, if one were to select all children diagnosed with cancer from a particular hospital, the proper control group would be those children who would have gone to that hospital had they developed cancer. The difficulty of defining this group will vary with the complexity of hospital referral patterns, which is especially true for childhood cancer because individual hospitals may have expertise related to specific diagnoses. In this commentary, we briefly examine the potential control group options for studies of childhood cancer etiology.

RANDOM DIGIT DIALING

In recent decades, random digit dialing (RDD) was the method of choice for many studies, both population based and hospital based (4–6). The Waksberg method (or some variation thereof) was typically used (7, 8). In population-based studies, it was reasonable to assume that controls in a defined geographic area who had telephones would have been ascertained as cases (with telephones) had they developed the disease. Most case-control studies of childhood cancer were conducted through institutions affiliated with the former Children’s Cancer Group; these institutions were estimated to treat approximately 50 percent of all children diagnosed with cancer in the United States (9). Since cancer treatment was centralized at a fairly small number of hospitals affiliated with the Children’s Cancer Group within a given geographic area, the study design assumed that the identified controls would have been treated at one of the Children’s Cancer Group hospitals had they developed the disease.

Several of the shortcomings associated with RDD (including lower response rates, possible selection for people who stay at home, noncoverage of people who do not have a telephone, and possible bias with respect to number of siblings) were considered acceptable given the centralized national sampling frame and a lower cost than required for other methods (10, 11). Until the mid-1990s, RDD worked well: in one case-control study of childhood brain tumors conducted during 1987–1991, 89 percent of households were screened successfully, and, for 73 percent of those found eligible, the telephone interview was completed (12). However, with the increasing use of cellular telephones, answering machines, and caller identification, along with the persistent annoyance of telemarketers, large-scale use of RDD is becoming inefficient and potentially subject to major selection bias. Although declining response rates with RDD were not confirmed in one recent report that evaluated rates in several published studies (13), it has been our experience...
in conducting national case-control studies that RDD response rates are worse now than they were a decade ago. This observation, of course, does not preclude the fact that RDD may be the most appropriate method for certain situations or in select localized areas of the United States.

**ALTERNATIVE CONTROL GROUPS**

Given the limitations of using RDD nationally, other control groups, such as hospital controls, friend controls, neighborhood controls, school controls, and cousin controls, have been considered, but each group has limitations that affect feasibility and appropriateness (table 1). Thus, we were very encouraged by the work of Ma et al. (14) in this issue of the *Journal*. In this study, the investigators compared the feasibility and representativeness of birth certificate and friend controls in a case-control study of childhood acute lymphoblastic leukemia in Northern and Central California. Birth certificate controls were obtained from the California Birth Registry, while friend controls were nominated by a family member of the acute lymphoblastic leukemia case. Using a number of selected variables, including parental age, number of livebirths, child’s birth weight, pregnancy losses, and time since last birth, Ma et al. reported that birth certificate controls who participated in their study were quite similar to a random sample of “ideal” birth certificate controls selected without tracing. Furthermore, they found that the friend controls were significantly different from the ideal population controls, further emphasizing the potential for selection bias when this control group is used (15). Although the results of this study are encouraging, some issues need to be considered further.

### Representativeness

The investigators (14) acknowledge that they achieved an overall response rate of 49 percent, which is similar to that in many recent studies using RDD. The actual response may have been lower or higher because they assumed that the proportion of eligibles among those who initially refused to participate and persons unable to be located was the same as that among those they were able to contact and screen for eligibility. However, they reported that since several of the variables recorded on the birth certificates of the ideal controls and the participating controls were similar, the participating controls likely provided a representative sample. Although it is understandable that the investigators chose variables that are more accurately recorded on birth certificates, it will be of interest to see how comparable these groups are with respect to variables that are not as consistently recorded but are also of etiologic significance, including maternal cigarette smoking, alcohol consumption, and method of delivery, for example.

### Feasibility in tracing

Some concern exists that use of birth certificates for case-control studies involving older children might be more difficult because of population movement. Ma et al. (14) provided reassuring data that the average number of searches required for older children was only marginally higher than that for younger children. However, given that childhood acute lymphoblastic leukemia affects children mostly under the age of 6 years, it will be of interest to see how well this control selection method works for a pediatric disease that affects primarily older children and adolescents.

### Potential for expansion nationwide

Finally, this study (14) raises the possibility of using birth certificates for nationwide pediatric control selection. In particular, concerns about ineligibility due to out-of-state residency would be mitigated if nationwide birth certificates were available. As Ma et al. (14) report, at least one other case-control study (in New York State) used birth certificates and found similar response rates (16). We are currently determining the feasibility, administrative and otherwise, of obtaining birth certificates from most (if not all) of the 50

---

**Table 1. Possible control groups for pediatric case-control studies**

<table>
<thead>
<tr>
<th>Control group</th>
<th>Strengths</th>
<th>Limitations</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random digit dialing</td>
<td>Population based</td>
<td>Increasing lack of participation; potential selection bias difficult to quantify</td>
<td>11</td>
</tr>
<tr>
<td>Hospital</td>
<td>Easily accessed; perhaps less subject to recall bias</td>
<td>Referral patterns often difficult to quantify; Berkson's bias possible</td>
<td>17–22</td>
</tr>
<tr>
<td>Neighborhood</td>
<td>Participation likely to be higher; easily described</td>
<td>Impractical on a nationwide basis</td>
<td>11, 23</td>
</tr>
<tr>
<td>School</td>
<td>Easily identified</td>
<td>Possible overmatching on exposure; applicable to studies of school-age children only; likely to be administratively difficult on a nationwide basis</td>
<td></td>
</tr>
<tr>
<td>Family/cousin</td>
<td>Easily identified; high participation</td>
<td>Possible overmatching on exposure; difficult to study genetic effects</td>
<td>24, 25</td>
</tr>
<tr>
<td>Friend</td>
<td>Easily identified; high participation</td>
<td>Overselection of gregarious persons; possible overmatching on exposure</td>
<td>26–28</td>
</tr>
<tr>
<td>Birth certificate</td>
<td>Population based; availability of comparison variables for nonparticipants; easy to match on birth certificate variables</td>
<td>May be difficult to find and enroll subjects; participation may vary by geographic location</td>
<td>14, 29</td>
</tr>
</tbody>
</table>
states (as well as the District of Columbia and New York City) for control selection. Notably, the Children’s Cancer Group recently became part of the newly formed Children’s Oncology Group, a group of over 200 hospitals and institutions that treat the vast majority of children with cancer in the United States (9). Since the Children’s Oncology Group now forms a nearly population-based sample of childhood cancer, there are plans to develop a national pediatric cancer registry in collaboration with state and regional cancer registries. Because RDD is likely no longer a viable option for control selection in future studies that would use this registry, identification of alternative control groups is of utmost importance. The work by Ma et al. provides evidence that birth certificates offer a potentially feasible and representative control group.

ACKNOWLEDGMENTS

This work was supported by National Cancer Institute grants U01-CA98543, R01-CA79940, and R01-CA75169 and the Children’s Cancer Research Fund.

Drs. Ross, Olshan, and Bunin are members of the Children’s Oncology Group Epidemiology Steering Committee.

REFERENCES