Introduction

Epidemiology and biostatistics are two of the foundations of public health science and practice (Institute of Medicine, 1988). Graduate and undergraduate degrees in schools of public health all include some level of training in these basic disciplines. As an epidemiologist, I consider the training to be vital to provide future researchers and practitioners of public health the analytic reasoning and interpretation skills to (1) understand and interpret research publications, (2) plan and execute research and evaluation studies, (3) provide scientific information about causal evidence, and (4) use information about risk factors and cause to shape programs and policies. For public health practitioners who are epidemiology specialists, the training and the skill set is more complex and expansive. When teaching the complexity of analytic reasoning in epidemiology, we often search for relevant, real world examples that permit us to impart both methods and broader public health content. Disability epidemiology provides a rich opportunity on both counts. It offers intellectually stimulating examples on issues that span the disciplines of both epidemiology and biostatistics. These issues, some of which are discussed in this chapter, relate to everything from pragmatic problems in field research to theoretical frameworks about the impact of risk factors across multiple levels of personal and environmental influences. Special populations, including people with disability (PWD), also offer an opportunity to examine and grapple with the public health challenge of disparities, and through that challenge, to help the disciplines of epidemiology and biostatistics to evolve. Unfortunately, there is an uneasy and even disrespectful relationship between the disability world and the world of public health, including epidemiology, when such an interface exists at all.

E.M. Andresen (*)
Department of Epidemiology and Biostatistics, College of Public Health and Health Professions, University of Florida Health Sciences Center, PO Box 100231, Gainesville, FL 32610, USA
e-mail: andresen@phhp.ufl.edu
This chapter provides examples of the opportunities and limitations of the discipline of epidemiology as a tool for research and for teaching about disability. Biostatistical techniques are discussed as they facilitate or impede disability research, but this chapter does not purport to offer an expansive overview of statistical methods. The chapter is intended for public health students or practitioners with some basic epidemiology training, and therefore it does not attempt to explain basic epidemiology and biostatistics, except where relevant to specific examples related to etiology and outcomes research. There are excellent epidemiology methods textbooks that cover traditional methods and analysis issues in substantially more depth (e.g., Gordis, 2000; Koepsell & Weiss, 2003; Rothman, Greenland, & Lash, 2008; Szklo & Nieto, 2006).

For epidemiologists, I hope to evoke interest in a new sub-discipline. Epidemiology, like other disciplines, has its “lumpers” and “splitters.” Is someone trained in epidemiology just an epidemiologist? Or should they always apply an adjective to explain their particular area of interest, expertise, and research? Since most meetings of epidemiologists include proceedings grouped by specialty topics (cancer, genetics, maternal and child health, etc.: see the agenda for the Society for Epidemiologic Research at http://www.epiresearch.org/meeting/) it is likely that disability epidemiology has some small group of professionals who cluster within the definition, as is also the case of public health overall as described below. Like any area of science, disability epidemiology requires – and does not yet have – spirited debate about its contents, breadth, rigor, and nature. As discussed later in this chapter, the concepts and content of disability epidemiology derive naturally from a number of other epidemiology disciplines, like injury, maternal and child health, aging, and outcomes research. Despite some differences in methods and focus, neither biostatistical nor epidemiology principles alter remarkably as one moves from one content area to another. In this spirit, I offer examples that can be shared in introductory and methodology courses in epidemiology. Rather than limit consideration of disability to special courses, it would strengthen the training of epidemiologists to have public health and research examples on disability topics in general coursework. For students, there is nothing more limiting than learning epidemiology from a methodologist with a very narrow content focus. For faculty, however, examples outside of one’s own area of expertise are hard to find and time-consuming to compose. In addition to methodological examples in this chapter, there are a number of potential practice problems that I hope instructors will use as part of training examples, homework, or exams and quizzes.

There is a substantial challenge in linking traditional epidemiology and biostatistics and their methods with disability research. A few of these issues, such as field research methods and interviewing, or drawbacks to classifications and definitions of disability, are taken up in subsequent sections, and include such basic issues as defining disability and national surveillance for disability and its correlates. But the challenges are not merely methodological. Challenges also relate to core public health epidemiology concepts like the definition of prevention, and the active participation of “subjects” in the conduct of research.
Public health has long explained itself as a discipline aimed at finding causes of disease, and identifying interventions to prevent disease, impairment, and disability. Epidemiology is the discipline that provides the “how” of such investigations and biostatistics provides tools to determine, and report the statistical “significance” of the research. The heavy placement on primary prevention as the goal of these efforts has often translated, explicitly or implicitly, into a goal of “preventing people” with diseases, impairments, or disabilities. Epidemiologists need to challenge this view, expanding their definition of etiology to include PWD as another population in which we have an interest in improving outcomes across the full range of health risks (Andresen, 2004). Health events as diverse as influenza and breast cancer are relevant to PWD. Most importantly, epidemiologists need to embrace research about the outcomes of impairment and disability. Spinal cord injuries occur, and people with these injuries lead productive lives. The determinants of successful outcomes – people fully integrated into their chosen lifestyles – are also a proper arena for epidemiologic research and biostatistical methods. However, we cannot expect nor do we want a retraction of the research efforts into understanding and preventing events like spina bifida from nutritional deficits, or of spinal cord injuries that result from traffic crashes, or examinations into the causes of strokes or hearing loss. The point is to state clearly that prevention does not have to mean prevention of people; but the process of making personal or societal changes that can change the incidence or severity of impairments.

There is another aspect of disability epidemiology that will cause concern to some public health researchers: participation of subjects and their communities in the design, conduct, and interpretation of research. There certainly are proponents of the notion that epidemiologic scientists should draw from and inform the populations they study (American College of Epidemiology, 2000; Hagey, 1997; Higginson & Chu, 1991; Koplan, Thacker, & Lezin, 1999; Sandman, 1991) and the debates about epidemiologists’ ethical responsibilities have been numerous (e.g., Coughlin & Beauchamp, 1996; Feinleib, 1991; Gordis, 1991; Stoto et al., 2001; Zahm, 1992). Historically, the prevailing view was that the proper conduct of research requires that participants be uninformed about the nature of the investigation to avoid contaminating their information or response. In a randomized trial, for example, the “blinding” of study participants to their treatment is considered evidence of the best science (Chalmers, Celano, Sacks, & Smith, 1983; Meinert & Tonascia, 1986; Pocock, 1983). Consider, then the disability community’s expectation that participants will be involved in the design, conduct, and interpretation of research: “Nothing about us without us” (Charlton, 1998). People with disability expect participatory action research (e.g., see Andresen, 2004; Delman, 2007; Eckhardt & Anastan, 2006; Gallagher & Scott, 1997; Marshall, Kendall, Catalano, & Barnett, 2008; Minkler et al., 2002; Mirza, Anandan, Madnick, & Hammel, 2006; Morrow, Arunkumar, Pearce, & Dawson, 2007; Seekins, Smith & McCleary, 1992; Stewart & Bhagwanjee, 1999; Van Niekerk, Lorenzo, & Mdlokolo, 2006; White, Suchowierska, & Campbell, 2004), but these expectations have been viewed as anathema to traditional research methods. Nonetheless, the mandate is clear from the disability community (e.g., Rioux, 1997;
Stone & Priestly, 1996), some funders (National Institute of Disability and Rehabilitation Research, 1998), and more recently, from the newest recommendations on the content and methods for public health (Institute of Medicine, 2003). While examples of participatory action research in epidemiology are sparse (Hagey, 1997; and see Austin, 2002, for a notable example of genetic epidemiology involving a community), examples in public health include work with communities on HIV/AIDS (Roy & Cain, 2001; Stajduhar, Lindsay, & McGuinness, 2002), occupational health (Keith et al., 2001; Rosecrance & Cook, 2000), access to care (Reese, Ahern, Nair, O’Faire, & Warren, 1999), health promotion and injury prevention (Gallagher & Scott, 1997; Ravesloot et al., 2007; White, Mathews, & Fawcett, 1989), and minority groups (Marshall et al., 2008; Morrow et al., 2007; Reese et al., 1999; Schulz et al., 1998).

Good questions may require substantially better theoretical foundations – a foundation lacking in much epidemiological research – and theories and models of the disablement process will need to learn from other disciplines and opinions. Fortunately, we are urged, as a discipline, “to widen the scope of epidemiology; advance knowledge, science and the ability to protect the health of the public” (Beauchamp et al., 1991). Surely this can involve a more thorough incorporation of the communities we “study” such as people with disability, into our research teams. An increasingly important segment of clinical epidemiology seeks to define and incorporate patient centered or patient generated outcomes into research questions by including patients with experience in planning and conducting studies (e.g., Fowler et al., 1999; deLateur, 1997; Fritz, Delitto, Welch, & Erhard, 1998; Curtis, 1998; Gilliam et al., 1997). The notion that research participants have important ideas to “add to the pot” is therefore not completely foreign to research (Clancy & Andresen, 2002; Hagey, 1997). We revisit the issues of research participation and ethics and involving the community in epidemiologic research at the end of this chapter. See also Chapter 6 “Disability and Health Inequity” that discusses disability and ethics in detail.

Defining and Some Parameters of Disability Epidemiology

Epidemiology has expanded its scope beyond the narrow focus on disease or illness to look at health events, health states, and health differences (Porta, 2008). Although disability epidemiology has not been defined explicitly, the scope of general epidemiology would suggest that the definition should certainly include study of basic elements of descriptive epidemiology (who, what, where), etiological determinants of impairments, and the frequency and predictors of different outcomes of disability. All of these questions incorporate the common epidemiology study designs and methods. Questions that ask students to consider study designs using disability topics are included in the Appendix. Additional descriptive disability epidemiology examples also are included in the section on surveillance and secondary data sources.
Importantly, disability must be defined as a state that is largely independent of health and health status. The distinction between the two is difficult to incorporate into traditional epidemiology, however. While epidemiologists are comfortable with the notion of prevention occurring at primary, secondary, and tertiary levels, in disability research, prevention includes both prevention of new primary disabilities or impairments, and prevention of the adverse outcomes of disability. In addition, as described by the World Health Organization model of disability below (the ICF), disability is a state that is inexorably connected to the environment in which people live. That is, with a properly functioning social, built, and policy environment, a person with disability would also function well, and would enjoy improved health outcomes. The separation of health and disability is accomplished in the *Healthy People 2010* chapter on disability (Centers for Disease Control and Prevention, 2001a). Prominent in this document is the following basic assumption about disability:

Disability is a demographic descriptor rather than a health outcome. It should be used to monitor disparities in health outcomes and social participation. The Americans with Disabilities Act (ADA) provides an important rationale for universal collection of disability status [in data collection].

Defining disease, health outcomes, health states, or health events for purposes of disability epidemiology are further challenges. The temporal nature of these events or states is important. A disease or health outcome may be an *event* that occurs at a specific point in time. For example, an automotive crash that results in injury, a suicide attempt, or a birth injury would be classified as events. A disease also may be a *state* that characterizes an individual during some time period. That is, it has duration. For these conditions or diseases, the particular time of onset (or diagnosis) is often of interest. Examples include depression, Parkinson’s disease, disability, or a particular level of health-related quality-of-life (HRQoL). A disease or health state also has a course that helps to define how it is dealt with in epidemiologic studies.

Some diseases occur only once and the individual does not (given current medical understanding) recover, e.g., Alzheimer’s disease and Parkinson’s disease. However, the *impact* of the state, disease, or characteristic is not necessarily static and may change over time. For disability epidemiology, this dynamic state fits well into the concepts of the International Classification of Functioning, Disability and Health (ICF: World Health Organization [WHO], 2001a). For a person with a spinal cord injury (SCI), there may be no change in the biological description of their injury level and neurologically defined impairment as time passes. But there might be substantial variation over the course of his or her lifetime in the impact of SCI, especially as impairments are affected by environmental variables. Other conditions can affect an individual multiple times, e.g., upper respiratory tract infections, depression, or myocardial infarction (MI or heart attack). A disease or health condition can be secondary to some primary condition, for example if a person has a spinal cord injury, our interest may be in the occurrence of secondary conditions and health events, such as pressure sores, urinary tract infections, etc.
In epidemiology, the course of a disease or condition is important because the measures of excess risk from an exposure or experience require some understanding of a person’s or population’s risk status. An individual is either “at risk” or “not at risk” for a given disease or event at a given time. “At risk” does not necessarily imply “at high risk” relative to others: it means that there is a nonzero chance of developing a disease. “Not at risk” means an individual is considered to have a zero chance of developing the disease. So, for example, a person with a spinal cord injury (SCI) may be at risk for a urinary tract infection (UTI) and this risk is higher, on average, than that of a person who is fully ambulatory; but the latter is also “at risk,” although the risk is much smaller. Different risks based on individual characteristics are what we measure in epidemiology to identify risky exposures and define groups with need of specific intervention(s). An individual’s “at risk” status may change over time. In the example above (urinary tract infection), it is particularly important to discover if there are the intervenable risk factors that place some people with spinal cord injury at special risk for an adverse outcome; or in other words, discover what risk factors there are in addition to the primary impairment of SCI. Added to this mix of factors that describe a person’s possibly elevated “at risk” status for disability is the need to measure and understanding environment. For example, having a personal care assistant may have a large impact of the occurrence of pressure sores for a person with SCI, and good transportation and accessible food stores contribute to the nutritional status of a person with is blind. Some variables that shape the association between SCI and UTI may not turn out to be ones on which we can intervene, but may give us insight into groups with extra risk, for example differences in risk of UTI according to age, gender, and other health conditions.

In summary, disability epidemiology includes the usual aspects of epidemiologic investigations. These include descriptive and analytic epidemiology, and the full spectrum of events and health states that arise in all epidemiologic work. The examples above help refine a few special issues in disability study designs. Additional methodological issues are further elaborated with examples about measurement and classification later in this chapter, and potential questions for training in epidemiology on study design issues are included in the Appendix.

**Contributing Epidemiology Disciplines**

As noted earlier, disability epidemiology is a loosely defined discipline. There are no societies of similar scientists nor are there scientific epidemiology journals where one would examine the ongoing work of the discipline (in contrast to other branches of epidemiology that produce such journals as *Neuroepidemiology; Epidemiology & Infection; Paediatric & Perinatal Epidemiology; Cancer Epidemiology, Biomarkers & Prevention; and Social Psychiatry & Psychiatric Epidemiology*). However, there is a newly developed journal for the broader aspects of public health with substantial numbers of epidemiology
articles, *Disability & Health Journal* (http://www.elsevier.com/wps/find/journal
description.cws_home/713446/description). Disability epidemiology derives
much of its content and methodology from traditional epidemiology and other
well established disciplines. Content areas tend to be defined either by a medical-
model for classifying diseases, or by methodological discipline. For example,
injury epidemiologists and epidemiologists working in content areas like
gerontology and neurological diseases contribute to the issues often dealt with
in disability epidemiology, such as challenges in collecting information from
respondents with cognitive or communications impairments. Since disability
research seeks to expand past a medical model (e.g., Institute of Medicine,
1991, 1997, 2005, 2007; World Health Organization, 2001a), we need to be
cautious about defining the discipline along these same models. However, the
textbooks and reference examples that we use in training epidemiologists also
are grouped along these same basic disciplines and models (e.g., Checkoway,
Crawford-Brown, & Pearce, 1989; Lawton & Herzog, 1989; Pless, 1994;
Wallace & Woolson, 1992). At present, we can embrace the contributing disci-
plines and methods, observe and respect their boundaries; but work toward
understanding the commonality of each in defining a new conglomerate
discipline.

**Disability Surveillance Data and Secondary Analysis**

**Introduction**

Many public health questions and program plans regarding disability are based on
ongoing surveillance systems. This section provides an overview of some of the
sources for the data regarding disability, and their use and problems. As a nation,
we need to know various characteristics of our population in order to monitor our
progress and allocate resources. We may have a relatively simple question, “How
many people are there in the U.S.A. with disabilities?” but the answer depends on
(1) our definition of disability, and (2) the mechanism we choose for collecting the
information. Epidemiologists are generally comfortable that definitions and clas-
sifications may be different depending on a specific question or need. However, the
simple problem of counting people with disabilities makes it problematic to have
many definitions. The report generated from the Chapter 6 “Disability and Health
Inequity” (Centers for Disease Control and Prevention, 2001a) specifically recom-
mends consistent definitions, universal collection of disability status, and method-
ological studies to help provide more accurate and useful data (see Fig. 1 for
recommended universal surveillance questions). In fact, a variety of surveillance
sources are used to answer questions about disability prevalence and these are the
focus for this section. In addition, the general topic of sources for secondary data
analysis is covered here and in the Appendix. Challenges in obtaining disability
statistics are addressed in *Healthy People 2010* (Chapter 6 “Disability and Health
Inequity”). However, a number of surveillance efforts that currently exist do provide useful information. For a large number of publications that include national disability statistics, see the web site of the National Institute for Disability and Rehabilitation Statistics and Training Center (http://dsc.ucsf.edu/UCSF/). In addition, there is now a United Nations effort collaborating on disability statistics and promoting uniform reporting of disability statistics, called “The Washington Group” (http://unstats.un.org/unsd/methods/citygroup/washington.htm).

Federal and state surveillance systems offer data that can be used for preliminary analytic work on a broad array of disability and health topics. Entire careers can be built around analysis of these data. For many epidemiologists, the complex sampling designs that produced these data sets and the resulting sampling weight and statistical variance complexity may require additional statistical training or working with a survey statistician (e.g., see Kneipp & Yarandi, 2002). The quality of data and uses and limitations of such systems also are described on their web sites (see also Best, 1999; Bosco, 2001; also, for issues in data quality for administrative data, see Iezzoni, 2002). As the examples below demonstrate, most of these data resources have yet to be fully analyzed in relation to disability epidemiology and health outcomes associated with disability.

Measures and Classification of Disability

Surveillance Measures: What question captures the definition we want for surveillance purposes? As described by Adler (1996), there are many choices to address the question (see also Institute of Medicine, 1991, page 49 for an example of multiple federal estimates based on different questions on functional limitations). Interestingly, surveillance classification questions carry implications that can clash with disability advocacy and politics. A broadly defined question will add to the population size that is claimed for the advocacy community, giving the issues of concern to this community greater social weight. Yet many of those counted by a broadly applicable question will not view themselves as having a disability, nor will

<table>
<thead>
<tr>
<th>Question</th>
<th>Response Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Are you limited in any way in any activities because of physical, mental, or emotional problems?</td>
<td>(Yes, No, Don’t know/not sure, refused)</td>
</tr>
<tr>
<td>2. Do you have any health problem that requires you to use special equipment, such as a cane, a wheelchair, a special bed, or a special telephone?</td>
<td>(Yes, No, Don’t know/not sure, refused)</td>
</tr>
</tbody>
</table>

These questions have been used in the National Health Interview Survey, the National Health & Nutrition Examination Survey, and the Behavioral Risk Factor Surveillance System.

Fig. 1 Healthy people 2010 “universal” disability questions for surveillance

1 These questions have been used in the National Health Interview Survey, the National Health & Nutrition Examination Survey, and the Behavioral Risk Factor Surveillance System.
they be viewed as others as such. Their inclusion in a count could have the effect of diluting the force with which disability concerns are expressed. Questions that are worded narrowly, however, may exclude people who are seen and see themselves as part of the disability numerator. Decisions about the choice of questions that provide the data for public health need to reflect about the exact purpose of the surveillance. In general, surveillance measurements opt for a very broad and inclusive classification system. A few examples help demonstrate these issues.

Table 1 shows how one of the disability questions works (Nanda & Andresen, 1998) that has been used by the CDC’s Behavioral Risk Factor Surveillance System: “Are you limited in any way…” (e.g., Andresen et al., 1999b). If we wished to “count” people who had a traumatic brain injury (TBI) the fact that only 50% of them said “yes” to having a limitation would be a problem. On the other hand, the fact that only 50% of people with TBI reported limitations is a good sign: they distinguish between their disability and the negative connotation of “limitation.” However, for politics and program planning, an accurate and clear “count” is needed. The Appendix includes questions that exemplify the effect of using broad or narrow definitions of disability/mobility impairment on the sensitivity and specificity of Behavioral Risk Factor Surveillance System (BRFSS) data.

In another example, the combination of questions on limitations and use of any of a long list of special equipment (used on the 1998 and 1999 BRFSS) produced disability prevalence estimates ranging from 13.9% (South Carolina) to 18.8% (Alabama) among 11 states that participated in the CDC’s special BRFSS surveillance activities (Centers for Disease Control and Prevention, 2000). Similarly, the national prevalence of disability can be estimated using one of both of the 2010 universal disability questions, as shown in Table 2. In the 2001 BRFSS, a total of 41,262 respondents said they were either limited, used equipment, or both (19.8%) – people

Table 1 Percentage of persons who indicated they were limited in any way in any activity in a study of PWD (Nanda & Andresen, 1998)

<table>
<thead>
<tr>
<th>Group</th>
<th>Total interviewed</th>
<th>% who answered “yes” to screening question on limitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traumatic brain injury</td>
<td>16</td>
<td>50.0</td>
</tr>
<tr>
<td>Spinal cord injury</td>
<td>195</td>
<td>91.3</td>
</tr>
<tr>
<td>Parkinson’s disease</td>
<td>26</td>
<td>80.8</td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>16</td>
<td>75.0</td>
</tr>
<tr>
<td>Nursing home residents</td>
<td>44</td>
<td>66.0</td>
</tr>
<tr>
<td>Assisted living settings</td>
<td>212</td>
<td>50.0</td>
</tr>
</tbody>
</table>

Table 2 Prevalence* of disability based on two surveillance questions from the behavioral risk factor surveillance system

<table>
<thead>
<tr>
<th>Question/classification</th>
<th>Uses equipment</th>
<th>No</th>
<th>Total persons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity limitations</td>
<td>Yes</td>
<td>10,045(4.8%)</td>
<td>28,118</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>3,099</td>
<td>167,273</td>
</tr>
<tr>
<td>Total persons</td>
<td></td>
<td>13,144(6.3%)</td>
<td>195,391</td>
</tr>
</tbody>
</table>

*Unweighted prevalence
who met definitions based on both questions, however, comprised only 4.8% of the population (data available from http://www.cdc.gov/brfss/ti-docs.htm).

**Detailed Classification Systems – the ICF.** The current theoretical model of disability is multidimensional, incorporating individual and environmental factors (World Health Organization [WHO], 2001a). The World Health Organization’s *International Classification of Functioning, Disability and Health* (ICF) provides a framework and a coding scheme that captures information on domains of body functions (e.g., hearing functions; code b210), body structures (e.g., structure of inner ear, cochlea; code s2600) activities and participation (e.g., conversation; code d350), and environment (e.g., sound; code e250: individual attitudes of strangers; code 2445). In its full application, the measurement is complex and not readily applied to large-scale surveillance surveys, especially if the purpose of the surveillance system is to provide information to a broad spectrum of public health practitioners. However, future development of surveillance tools that capture one or more aspect of the ICF may be developed. Presently, the BRFSS provides substantial data on disability, but does not incorporate the ICF: as described below, in-depth aspects of surveillance measures of disability have been quite limited in this telephone survey. The ICF provides a set of codes for minimal health systems or surveys (World Health Organization, 2001a, page 253) that are somewhat daunting even in considering their relative brevity. However, the ICF was recommended, at least in principle, as the classification system for coding functional status in medical encounters data by the *National Committee on Vital and Health Statistics* (Department of Health and Human Services, 2001). In addition, the ICF is proposed as the unifying framework for disability measurement in public health (Lollar, 2002; Lollar & Crews, 2003).

Questionnaires that are substantially longer then usual surveillance tools have been developed for classifying people according to the ICF. The WHO has tested a comprehensive tool for measuring disability and the environment, the WHODAS II (World Health Organization, 2001b). This instrument provides a profile of functioning across six activity domains, as well as a general disability score. Published uses and testing of this instrument are limited (e.g., see Cieza et al., 2002; van Tubergen et al., 2003). Application of this survey is primarily clinical, but in addition to the standard 36-item version, there also is a 12-item version (which reduces administrative time from 20 to about 5 min) and future iterations and uses of the WHODAS may contribute to surveillance of disability (see World Health Organization, 2001b). To date, applications of the WHODAS II are very limited in the United States (e.g., see Chisolm, Abrams, McArdle, Wilson, & Doyle, 2005; Chwastiak & Von Korff, 2003; McArdle, Chisolm, Abrams, Wilson, & Doyle, 2005), although it is used more frequently in Canada and Western Europe (e.g., Luciano et al., 2010; Wang, Adair, & Patten, 2006). Other applied survey methods which use the ICF to code results have been developed, including the Craig Hospital Instrument of Environmental Factors© (for information contact dmellick@craig-hospital.org), which was used for the 1999 BRFSS in Colorado. In general, however, given the complexity and training issues for even clinical applications, the research applications of the full ICF are likely to remain in clinical rehabilitation and outcomes arenas, and not surveillance surveys (e.g., see Grimby & Smedby, 2001; Reed et al., 2008).
The Behavioral Risk Factor Surveillance System (BRFSS). The Behavioral Risk Factor Surveillance System (BRFSS) is an annual survey conducted in all American states and territories (Gentry et al., 1985; Remington et al., 1988). Currently, it is the primary source for data regarding disability for states, but also contributes to national estimates and analysis of disability. The BRFSS is supported and supervised by the Centers for Disease Control and Prevention (CDC) although states may use their own funds to add optional sections to the survey or increase sample size among particular population groups. The BRFSS is a random digit-dialed telephone survey that covers health behaviors, medical diagnoses, demographic variables, and health-related quality of life (HRQoL). The HRQoL questions also have been used to provide estimates of disability, as noted above. There is a four-question core HRQoL module, Healthy Days, used by all states, and an optional ten-question module that has been used by some states (e.g., about half of states used it in 2000 or 2001). Questions include categorical questions (e.g., general health, health limitations) and questions using the last 30 days as a referent (e.g., poor physical health days, pain days). In 2001 and up to the present, the core BRFSS data for all states, DC, and territories included two questions on disability (see Fig. 1). With these additions, which permitted identification of people with disabilities and ascertainment of information about their specific impairments and their general satisfaction, the BRFSS became a valuable source of national and state level information about disability.

Core questions for the 2000 BRFSS included items measuring caregiving provided to older adults with disability; 15.6% of respondents reported they provided caregiving to someone aged 60 or older (McKune, Andresen, Zhang, & Neugaard, 2006). In addition, 28 states added a supplemental module on caregiving to the BRFSS in 2000 and/or 2001 (Fig. 2). This provided new information on the BRFSS respondents who said they needed either personal care assistance or assistance with routine needs, and how adequate the caregiving assistance was (Jamoom, Andresen, Neugaard, & McKune, 2008). In more recent years, caregiving has been promoted as a public health issue (Talley & Crews, 2007) and some states now report on the prevalence and correlates of caregiving (Bouldin, Akhtar, Brumback, & Andresen, 2009; McGuire, Bouldin, Andresen, & Anderson, 2010; Neugaard, Andresen, DeFries, Talley, & Crews, 2007).

The BRFSS offers remarkable opportunities for epidemiologic questions and analytic practice for graduate students. These might include questions used in the context of a specific health behavior (physical activity, smoking), specific diagnoses (e.g., arthritis, diabetes), or examine disparities in health care access. The Appendix provides a number of epidemiology methodology practice questions using BRFSS data and questions.

The National Center for Health Statistics (NCHS) National Health Interview Survey. The National Center for Health Statistics (NCHS) conducts the annual National Health Interview Survey (NHIS). The NHIS generally includes questions on physical function and general health status, but during 1994 and 1995 it also included a supplement on disability (NHIS-D). NHIS-D data entail face-to-face interviews with 202,000 adults and children (using parent proxies for child subjects). Follow-up interviews collected data on 33,000 persons with disability. NHIS-D
findings provided an exceptional opportunity to examine health care utilization and access among people with disability, especially in comparison to the general population. Other disability-related topics addressed in the supplement were: chronic conditions, impairments, functional limitations, activities of daily living (ADLs, which are basic self care activities), instrumental activities of daily living (IADLs, which are more complex self care activities), the use of assistive technology, transportation, personal assistance services, labor force participation, workplace accommodations, and disability benefits. Many detailed reports have been produced to describe disability among children and youth (Druss et al., 2000; Druss & Rosenheck, 2000; Hogan, Msall, Rogers, & Avery, 1997; Newacheck et al., 1998), and specialty topics on adults, for example injury (Guerrero, Sniezek, & Sehgal, 1999), sensory impairment (Hoffman, Ishii, & Macturk, 1998), and personal assistant services (LaPlante, Harrington, & Kang, 2002).

NHIS-D data also have been used to examine disparities in health care. For example, women with functional limitations were found to have lower screening rates for Papanicoulaou tests and mammograms (Centers for Disease Control and Prevention, 1998; Iezzoni, McCarthy, Davis, Harris-David, & O'Day, 2001). Disparities in health insurance coverage for children with chronic conditions were reported from these data (Silver & Stein, 2001). Russell et al. reported on increasing use of assistive technology over the period ending with the 1994 NHIS (Russell, Hendershot, LeClere, Howie, & Adler, 1997). While it would be useful

| 1. Are you limited in any way in any activities because of physical, mental, or emotional problems? |
| 2. [If yes] Because of any impairment or health problem, do you need the help of other persons with your PERSONAL CARE needs, such as eating, bathing, dressing, or getting around the house? | Earlier you reported that due to your impairment you need some assistance from another person with your PERSONAL CARE needs... |
| 3. Who usually helps you with your personal care needs, such as eating, bathing, dressing, or getting around the house? | Is the assistance you receive to meet your personal care needs: (usually, sometimes, or rarely adequate)? |
| 4. [If yes] Because of any impairment or health problem, do you need the help of other persons in handling your ROUTINE NEED, such as everyday household chores, doing necessary business, shopping, or getting around for other purposes? | Earlier you reported that due to impairment you need some assistance from another person with ROUTINE needs... |
| 5. Who usually helps you with handling your routine needs, such as everyday household chores shopping, or getting around for other purposes? | Is the assistance you receive to meet your routine needs: (usually, sometimes, rarely adequate)? |

**Fig. 2** Summary of questions related to caregiving. An optional module for the Behavioral Risk Factor Surveillance System (2000, 2001)
for a new supplement to be fielded like the NHIS-D, NHIS surveys always include some data that can address disability and disability outcomes (http://www.cdc.gov/nchs/nhis.htm).

The Medical Expenditure Panel Study (MEPS). MEPS data collection is supported by the Agency for Healthcare Research and Quality (AHRQ). MEPS data provide a broad set of measurements of health status and diagnoses, health care use and expenditure, and issues about access to care (Cohen, 2002; Cohen et al., 1996–1997). MEPS data consist of four components. The Household Component (HC) collects data on a sample of families and individuals who comprise households drawn from a nationally representative subsample of households that participated in the prior year’s NCHS. The Medical Provider Component (MPC), which covers hospitals, physicians, and home health care providers, supplements information from respondents HC respondents. The Insurance Component (IC) consists of two subcomponents. Data from the household sample on the health insurance held by and offered to HC respondents is the first. The second component comprises data collected from a sample of business establishments and governments throughout the United States. MEPS also periodically includes a Nursing Home Component (NHC).

Disability analyses based on MEPS have examined ethnic disparities in function (Carrasquillo, Lantigua, & Shea, 2000), child disability and expenditures (Chan, Zhan, & Homer, 2002; Elixhauser et al., 2002; Hanson, 2001), health care issues in mental health disability (Druss & Rosenheck, 2000; Egede, Zheng, & Simpson, 2002; Goldberg, Seybolt, & Lehman, 2002; Olfson et al., 2002a; Olfson, Marcus, Druss, & Elinson, 2002b), complimentary medicine therapies used by PWD (Druss & Rosenheck, 2000), and home health care (Mullner, Jewell, & Mease, 1999). Work by Yelin et al. the value of these data as a basis for understanding access to care and health care costs (Yelin, Herrndorf, Trupin, & Sonneborn, 2001). They examined the 1996 MEPS data for persons aged 16–64 with musculoskeletal conditions and found that there was a striking difference in medical expenditure for those with health insurance (annual expenditure $3,249) compared to those without insurance (annual expenditure $793). They estimated that nationally, medical expenditures for those with musculoskeletal conditions were about $193 billion, or 2.5% of the Gross National Product for 1996. Future studies using MEPS data could explore issues related to disability that crosscut diagnostic categories and that represent definitions of disability in line with the ICF. MEPS data provide a unique opportunity to assess economic and other impacts associated with disability, especially because of the population perspective of the data that allow comparisons between PWD and others (see http://www.ahrq.gov/data/mepsix.htm).

The Census Bureau Resources. The United States Bureau of the Census provides data on disability drawn from three primary sources: the decennial census of population, the Survey of Income and Program Participation (SIPP), and the Current Population Survey (CPS). Two of these data resources are briefly described here (see also http://www.census.gov/).
The US Census, conducted at 10-year intervals, is not primarily a health surveillance tool, and yet it provides denominator data for use in population research, policy and political analysis, and resource allocations. In 1990 and 2000 (and projected for 2010), the long form of the census included questions on functional status and disability, permitting better estimation of disability prevalence. The Census 2000 questions addressed instrumental and basic activities of daily living (ADL/IADLs i.e., self-care, mobility), activity limitations, working, and sensory impairments (Andresen, Fitch, McLendon, & Meyers, 2000). Census data, even from the sample-based long form, are representative and applicable to very small areas – as small as neighborhoods. In fact, the small local area comparisons may need to rely on the census because of the difficulty of using state (BRFSS) and national (NCHS) data at smaller geographic regions (Andresen, Diehr, & Luke, 2004).

The Survey of Income and Program Participation (SIPP) is a multistage-stratified sample of the US civilian noninstitutionalized population. The SIPP is a continuous series of national panels, with sample size ranging from approximately 14,000–36,700 households. The duration of each panel ranges from 2.5 to 4 years. Interviews are conducted in-person and by telephone. All household members 15 years old and over are interviewed with proxy responses permitted when household members are not available. The SIPP content is built around a “core” of labor force, program participation, and income questions. SIPP data include detailed information on cash and noncash income and data on taxes, assets, liabilities, and participation in government programs. The survey can also include questions in “topical modules” to permit targeted analyses on issues such as disability. The 1996 adult SIPP panel was asked about disability during wave 11 (1996) yielding data that provided estimates of prevalence based on functional activity restrictions (Centers for Disease Control and Prevention, 2001b). About 22% of adults reported having a disability, with higher prevalence among women. As in other national surveys, arthritis/rheumatism was the diagnostic category associated with the greatest proportion of disability. These brief examples of Census Bureau data and publications hint at their utility in disability epidemiology. In particular, the data might be subjected to complex analyses beyond surveillance statistical reports of prevalence and descriptive epidemiology.

Selected Methodological Issues in Disability Research

Introduction

Disability research raises unique methodological issues. Four that are discussed in the section that follows are sample size, measurement, multi-level studies, and the process of field research. As often happens in relation to methodology decisions, there are tradeoffs among solutions. For example, methods to include representative PWD in research efforts often involve increased costs. Efforts to increase sample sizes may require a less costly (and less precise) measurement strategy. In order to increase samples for statistical precision, it makes sense to combine people with different impairments. However, epidemiology is usually concerned with precise
measurements (e.g., exposures and case-definitions) to estimate valid exposure-outcomes relationships of interest. These issues, and potential conflicts between solutions, are discussed in more detail below.

Sample Size

In epidemiology, if a condition is not relatively common, it is hard to conduct statistically useful research on it. The following examples based on hearing impairment/loss in children show why this is true. Hearing impairment among infants and young children is challenging to detect and surveillance statistics equally hard to estimate. The CDC’s programs of research on this topic include an active surveillance registry for the Metropolitan Atlanta area (MADDSP Centers for Disease Control and Prevention, 2004). An estimate of the prevalence of severe hearing impairment among children aged 3–10 has been estimated as 1.1 per 1,000 children (Van Naarden, Decoufle, & Caldwell, 1999) and the estimate is higher for African American male children (1.4/1,000). If we wished to study this problem in more detail and understand the contributors to higher risk of hearing impairment, it would require over 445,000 children to be sure that we were confident that there is a statistically significant difference between 1.4/1,000 and 1.1/1,000. That is, hearing impairment is not common, and it would take a very large sample of children to determine that it was more frequent among African Americans. But let’s say we wanted to look at a new suspected cause of hearing loss in children (some new pregnancy exposure, e.g., a new popular dietary supplement among pregnant women that about 10% of pregnant women use). Table 3 shows how many children with and without the exposure would be needed to detect a significant difference based on exposure status (i.e., conclude that the exposure caused hearing impairment in statistical terms). The measure of effect in this hypothetical cohort is the relative risk (RR). You can reproduce this sort of

<table>
<thead>
<tr>
<th>Exposure effect size</th>
<th>Total study sample sizeb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative risk (RR)</td>
<td></td>
</tr>
<tr>
<td>1.1</td>
<td>3,030,000</td>
</tr>
<tr>
<td>1.5</td>
<td>150,000</td>
</tr>
<tr>
<td>2.0</td>
<td>46,300</td>
</tr>
<tr>
<td>5.0</td>
<td>6,200</td>
</tr>
<tr>
<td>10.0</td>
<td>2,300</td>
</tr>
</tbody>
</table>

a$p=0.05$, power=0.80, equal number of exposed and unexposed; frequency of hearing impairment in the unexposed = 1.1/1,000 children/mothers
bSample size estimates in this calculation have been rounded since the numbers were so large.
exercise using *EpiInfo*, a statistical program available from the CDC (Centers for Disease Control and Prevention, 2003; Sunk, Sangan, & Arner, 2000). Clearly, this may be a cost-prohibitive study if the effect size were modest – for example, to answer the question on causation based on a RR of 1.5, the study would have to enroll over 150,000 children/mothers to follow up and ascertain outcomes when children are between 3 and 10 years of age.

In some instances, rare events can be studied more efficiently by case-control methods. Table 4 shows the sample size required for case-control studies with the same effect sizes (here the odds ratio approximation to the relative risk, or OR). In this example, we also could take advantage of including a higher control-to-case ratio since we know that hearing impairment is rare and enrolling cases (children/mothers) will be difficult. If we use a single control per case, we can accomplish the same study with a medium effect size of 1.5 as our cohort above with 1,914 subjects (and 957 cases). Increasing the control ratio to 3 per case, we would need a slightly larger total sample size, but only 618 cases of children with hearing impairment. The problems of case-finding, enrolling and interviewing subjects, ascertaining exposure with accuracy, and the potential for problems of recall bias might still make this a difficult study. And it is important to note that hearing impairment, while rare, is still a common event compared to other specific impairments. For example, about 40 per million Americans sustain (and survive) a spinal cord injury each year, making this a rare event that is difficult to study on the population level (National Spinal Cord Injury Statistical Center, 2001). These examples underscore the problem of conducting disability research, especially with the aim of ascertaining etiology of specific impairments.

### Measurement Issues (Classification and Bias)

As noted above, combining groups is one strategy to overcome the problem of small samples. For example, we might combine specific diagnoses into functional categories such as mobility, communication, or learning/cognition impairments.

<table>
<thead>
<tr>
<th>Exposure effect size (OR)</th>
<th>Total sample size (cases only)</th>
<th>Total sample size with 3:1 controls per case (cases only)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>37,400 (18,700)</td>
<td>49,500 (12,380)</td>
</tr>
<tr>
<td>1.5</td>
<td>1,914 (957)</td>
<td>2,472 (618)</td>
</tr>
<tr>
<td>2.0</td>
<td>614 (307)</td>
<td>776 (194)</td>
</tr>
<tr>
<td>5.0</td>
<td>78 (39)</td>
<td>116 (29)</td>
</tr>
<tr>
<td>10.0</td>
<td>44 (22)</td>
<td>52 (13)</td>
</tr>
</tbody>
</table>

*p* = 0.05, power = 0.80; base frequency of exposure in control children/mothers is set in this example at 10%. This example is based on an unmatched study design.

Sample size estimate in this row were rounded since the numbers were so large.
This is part of the rationale for the ICF classification system of the WHO. The trade-off here is that heterogeneity of exposures and outcomes may bias relationships toward the null (no effect). For a study on etiology, combining people with different impairments will obscure causes. Such varied diagnoses as Parkinson’s disease, multiple sclerosis, HIV/AIDS, and spinal cord injury call all lead to mobility impairments but entail very different etiology. Thus, grouping people with mobility impairments in a study of etiology would obscure the very different causal pathways that can produce this outcome. For a study on secondary conditions, or outcomes research, this grouping of impairments might be fine as long as proper consideration was paid to differences of other characteristics and risk factors, such as severity.

In epidemiology, we generally prefer homogenous measures of both exposures and outcomes precisely because blending across disparate groups can mask effects. A real difference (measured, for example, by a relative risk) may be missed because of fuzzy classification. A real excess risk may be underestimated. For example, in a study concerning the risk of secondary conditions among people with disabilities, the sample included people with mobility impairments, mild cognitive impairments, and mental health disability (Andresen, Fouts, Romeis, & Brownson, 1999a; Prysak, Andresen, & Meyers, 2000). Based on these data, an important current question about risk of secondary conditions is whether people with mobility conditions, and specifically people who use wheelchairs, are at higher risk of particular secondary conditions like pressure sores. If we include people who use wheelchairs for diverse reasons, we may mask the differential effects of wheelchair use among subgroups. Table 5 shows how this can work. If we include all 126 people in our sample of 239 who are “exposed” to wheelchair use, we do, indeed, see that they have an increased risk of having skin problems (pressure sores, skin ulcers, etc.). The gender and age adjusted risk is 4.3; this is a pretty strong effect (and the 95% confidence intervals exclude 1.0). But people can use a wheelchair for many reasons. People with spinal cord injury (SCI) use wheelchairs because of paralysis, which could entail a much higher risk. If we define the exposed group as the 83 people with SCI, we see that the risk increases to more than 10 times baseline. In fact, the homogeneous group masked the very high effect of wheelchair use among people with

<table>
<thead>
<tr>
<th>Exposure group</th>
<th>Group sample number</th>
<th>Skin problems Number</th>
<th>Adjusted odds ratioa</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>No wheelchair (reference)</td>
<td>113</td>
<td>9</td>
<td>8.0</td>
<td>–</td>
</tr>
<tr>
<td>All wheelchair users</td>
<td>126</td>
<td>39</td>
<td>31.0</td>
<td>4.3</td>
</tr>
<tr>
<td>People with spinal cord injury who use a wheelchair</td>
<td>83</td>
<td>30</td>
<td>36.1</td>
<td>10.0</td>
</tr>
</tbody>
</table>

aOdds ratios are adjusted for age and gender
SCI. Some studies may need to use more homogeneous groups rather than broad heterogeneous classifications to elucidate such exposure-outcome casual questions. In this example, we fortunately would conclude that there is risk for even the heterogeneous group; but clearly the causal inference one would draw is stronger for people with SCI.

**Multilevel Studies**

Epidemiologists, historically, tended to eschew research where associations are made between variables based on grouped data, that is, those that are “ecological” in design. In a purely ecological study, we compare people at an ecological level in terms of exposure and outcome. The problem of inferences about cause-and-effect relationships from grouped data is called *ecological fallacy*.

However, when exposures really do operate at an ecological level as do many that have a potential impact on PWD (e.g., especially environmental ones, such as disability awareness media campaigns, laws, severity of winter weather, wheelchair accessible public transportation), there is a strong argument for applying a combined ecological and individual level study design (e.g., see Von Korff, Koepsell, Curry, & Diehr, 1992). In fact, the IOM recommends combined –personal and environment – models of public health training and thinking as a high priority (Institute of Medicine, 2003), although this typically refers to understanding the social influences on health for racial and ethnic minority groups. Aside from a need for more complex statistical analysis (i.e., mixed variance estimates), there are relatively few barriers to multi-level studies in disability epidemiology. The ICF codes for specific aspects of environment as experienced by individuals. Measures and models that are truly multi-level can provide a very rich area of research to understanding variations in levels of participation of PWD: at present, this kind of work is in its infancy, and few disability-relevant ecological-level measures have been defined.

**Field Methods in Disability Epidemiology**

Allan Meyers used the phrase “enabling our research” to describe what is needed to support the health assessment of people with disability (Meyers & Andresen, 2000). He also advocated what he termed “universal design” in research methods and measures borrowing the phrase from architecture and interior design. These ideas reflect the observation that is often a challenge to include PWD in research and that they are often, therefore, excluded, from research. A study conducted by Barnett and Franks (1999) underscores just one aspect of this problem. They demonstrated that deaf people are substantially less likely to own a telephone, and are therefore excluded from participation in surveys conducted by telephone. Even deaf people with telephones use TTYs, which may also lead to exclusion.
In general, field research, meaning finding subjects and collecting data from them, entails specific disability-related challenges. These include, but are not limited to (1) communication (e.g., telephone survey methods), and (2) validity of questions from general surveys when applied to disability groups. Below is a descriptive list of how these issues affect research (Meyers & Andresen, 2000).

- Standard sampling methods do not reach the same proportions of PWD as people without disabilities.
- If they are included in a sample, some PWD may be excluded by standard modes of questionnaire administration and interviews; others may be unable to participate to a limited extent do not allow them to complete questionnaires and interviews as comprehensively as people without disabilities.
- Even when PWD are included in a sample and able to complete instruments, many questionnaires include questions and measures that are offensive or simply irrelevant to PWD.
- Enablement of disability research is an ethical, legal, and methodological imperative.

Ideally, there should be full participation of people with disabilities in epidemiologic research, whether they are the primary focus of a particular study or just one statistically identified subgroup in broader population research. Consider the problem of a tradition of excluding women and older adults from randomized trials of heart disease prevention and treatment therapies. The result has been a need to play “catch up” in order to provide meaningful information for women (e.g., see Bennett, 1993; Buring & Hennekens, 1994; Gurwitz, Col, & Avory, 1992). It does not require disability advocacy or a full interpretation of the ADA (Americans with Disabilities Act) to make this same point about people with disabilities. The full population perspective on the health of Americans has to include PWD to be complete, especially with the changing age structure that includes larger proportions of older adults. Further, the health of PWD of all ages is a public health priority, and inclusion of PWD in surveillance and research must be a component of the successful conduct of public health functions aimed at health objectives for PWD.

**Protection of Human Subjects in Disability Public Health Research**

In general, research that is inclusive of people with disability is not different in terms of human subjects’ issues than other research. However, Human Subjects Research Committees (IRBs) often incorporate a category of special concern for research participants with disability that does not distinguish aspects that might really require extra steps. This kind of broad categorization of protection reminds us of the decades of research that excluded women of child-bearing age so as to protect them and their possible fetuses. The unstated assumption seemed to be that women did not know how to protect themselves and could not participate in
informed consent and decide themselves if they wished to participate. While this is not likely to be the explicit reason for having PWD in a category of IRB concern, considering special circumstances remains important for public health researchers.

It is possible that persons of legal age may include individuals with cognitive or intellectual impairment, or limited English language fluency. Public health researchers should embrace self-determination as a core ethic for human subjects research, and acknowledge that legal assent/consent may require additional safeguards. Informed consent procedures may include brief cognitive screening to determine subjects capacity to understand and freely give consent (e.g., using the MMSE [Folstein, Folstein, & McHugh, 1975]), review of literacy and language using literacy screening tests as reviewed by the Institute of Medicine (Institute of Medicine, 2001) for example the REALM (Davis et al., 1993) or TOFHLA (Parker, Baker, Williams, & Nurss, 1995), and/or oral and Braille formats of consent procedures.

Research participants with mental health disability might also require added considerations. To approve research under the federal regulations, an IRB should determine that subject selection is equitable and be particularly cognizant of the special problems of research involving vulnerable populations including persons with mental health disability. There is no standard definition of who is vulnerable as a person with mental health disability and there are no additional safeguards codified in regulation. However, individual states may have statues that specify who can consent for incapacitated/ incompetent/ decisionally impaired adults.

Researchers should consider their IRBs as a partner in the protection of human subjects’ process, consult with them about their plans, and consider IRB feedback to determine if additional safeguards should be included in the study to protect the rights and welfare of research participants with disability. Pragmatically, experienced disability researchers know that some cross-education on research including PWD may be needed to facilitate the sense of partnership with IRB staff and committees (White, Klatt, Gard, Suchowierska, & Wyatt, 2005).

**Future Directions and the Role of Epidemiology**

This chapter has described classic epidemiology methods and their relevance and benefit to research involving disability populations, issues, and outcomes. A number of methodological difficulties are present: for example, epidemiologists may express concern over the heterogeneous nature of disability classification as an “event” or “state.” Another area of concern for epidemiologists is the evolving theoretical nature of the disabling process, which overlaps to the subject of classification. Until epidemiologists become familiar with the WHO’s ICF and practice stating their research questions in this new framework, it will be difficult to blend the theory and quantitative science of disability research. Epidemiologists make few apologies for their relative lack of theoretical training, but in this instance,
cooperative thinking is required to move ahead. The recent IOM report on training public health practitioners and researchers makes a strong recommendation about the need for transdisciplinary research (Institute of Medicine, 2003). Epidemiology is not the sole focus of this encouragement, and disability was not included as an example in the report. However, there is substantial integration between the emerging priorities and competencies for public health professionals, and the recommendations made in this chapter.

Finally, a general area of research that is somewhat new, and may make epidemiologists uneasy, is the notion of involving the research population in the research process. We also include commentary concerning the involvement of people with disabilities in the process of research (Andresen, Lollar & Meyers, 2000; Meyers & Andresen, 2000), and incorporating the important directive from disability research, “Nothing about us without us.” (Charlton, 1998; Seekins et al., 1992). But the field of epidemiology and its ethical and disciplinary goals have evolved, and the potential issues raised here are now incorporated into our discipline and to public health practice and research. An early set of epidemiology ethical guidelines (Beauchamp et al., 1991) included that epidemiologists should “… widen the scope of epidemiology; advance knowledge, science and the ability to protect the health of the public.” This could easily be interpreted, in part, as a directive to free up from strict disciplinary boundaries in methods, and to try new scientific and methodological tasks to improve research aims. In addition, this early guide was specific about the respect and care in regard to the subjects of research, and the communities who required the information. Going even further, a more recent set of guidelines (American College of Epidemiology, 2000) discuses the potential requirement of involving community representatives in our research – this is a very interesting addition. Here the language is cautionary “To the extent possible and whenever appropriate…” But this goes on to issue of public trust, and being drawn into problems of “unempowered communities.” Here, the consideration of “participatory” approach to research projects is explicit. Another more recent discussion in epidemiology, and stated explicitly here, is that advocacy may be undertaken by epidemiologists, but must not impair scientific objectivity. Lastly, these discipline-specific guides regarding the ethics of including communities and subjects’ perspectives in research now can be incorporated into the new imperative for community-based participatory research for scientific and ethical reasons (Institute of Medicine, 2003). The challenges of disability research that confront epidemiology also provide substantial impetus to expanding methodological innovation, extending traditional epidemiological thinking, and learning from our knowledgeable subjects and communities.

Epilogue

During my years of training epidemiologists and researchers and discussing how epidemiology views (or does not view) disability, I have encountered mostly right-minded and caring professionals and students who find the barriers that people with
disabilities face inexcusable. The belief in social justice reigns strong in public health communities. It is therefore surprising, and even shocking to many, that the basic concept of “prevention” might be taken as a problem in public health thinking (Andresen, 2004). Two kinds of examples help me make this issue concrete.

The first example comes from public health history of defining prevention: even today health departments include “disability prevention” as part of their names. Examples include the District of Columbia (2010), Santa Barbara (2010), and Fresno County, California (2010). The leadership of the Centers for Disease Control and Prevention uses current mainstream terminology in naming their public health division “Disability and Health.” A visit to their website will provide an education to students and practitioners alike as to the full public health and research agendas in disability (see http://www.cdc.gov/ncbddd/dh/default.htm).

The second example speaks to the common response of good people that “of course” they do not mean they seek to prevent people. I use an example from an article in the New York Times to make the point that PWD may not be paranoid about the use of the term “prevention” (McBryde Johnson, 2003). An exceptional woman, Harriet McBryde Johnson, shared her story about her experience in a debate with Princeton Professor Peter Singer. Professor Singer has advocated for the possibility of euthanizing babies who whose lives he deems to be cognitively impaired enough to not be “persons.” The cover magazine read “Should I have been killed at birth? The case for my life.” In public health we can all answer her question easily and with spirit (“No, Harriet! Your life has meaning and value.”). We simply have to think more deeply about what would provide for equity, inclusion, and participation of people with disability in public health and epidemiologic research.

Appendix: Incorporating Disability Examples into Epidemiology Coursework, Questions, and Examples

Prevalence, Odds Ratios, Stratified Analysis (Confounding or Effect Modification), Study Design, and Causal Inference

As a project for a disability epidemiology course, M.P.H. students Tori Vahle and Mary Gould analyzed data from special surveys in Missouri adapted from the Behavioral Risk Factor Surveillance System (BRFSS). These were random digit dialed telephone surveys conducted in six Missouri counties between 1995 and 1997. The sample consisted of a total of 3,343 adults: 1,380 from rural and 1,963 from nonrural areas. Disability was defined as “limited in any way in any activities because of any impairment or health problem.”

(a) Table 6 shows the crude results of the study. Is there an increased prevalence of disability in rural areas compared to urban areas?

Answer: The odds ratio for this table is 1.14 (95% confidence interval [CI] = 0.96, 1.34).
Table 6  Risk of disability by urban/rural residence

<table>
<thead>
<tr>
<th>Residence</th>
<th>Limited/disability?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural</td>
<td>Yes 344, No 1,036</td>
</tr>
<tr>
<td>Urban</td>
<td>Yes 444, No 1,519</td>
</tr>
</tbody>
</table>

Table 7  Risk of disability by residence and age

<table>
<thead>
<tr>
<th>Age group</th>
<th>Limited (disability)</th>
<th>Not Limited</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural</td>
<td>Urban</td>
<td>Rural</td>
</tr>
<tr>
<td>18–64</td>
<td>132, 314</td>
<td>840, 1,244</td>
</tr>
<tr>
<td>65–99</td>
<td>210, 128</td>
<td>190, 254</td>
</tr>
</tbody>
</table>

(b) The data were also stratified by age. Table 7 shows these data (note: age is missing for 31 subjects). Accounting for age, how does the odds ratio change compared to your answer based on Table 6 and answer a. above? Why?

**Answer:** The rural population is older than the urban population; we can see that 20% of the urban group is over age 65 and 29% of the rural population is over age 65. Since disability increases with age (due to increasing chronic conditions), the crude odds ratio is biased high just because of the older age of rural adults. The age-adjusted OR is 0.98 (95% CI = 0.83, 1.17). However, if one examines the stratum-specific results, it appears that the results are reversed for the two age groups. For younger adults, the effect of living in a rural area is protective (OR = 0.62, 95% CI 0.50, 0.78) but for older adults, it is associated with an increased risk (OR = 2.19, 95% CI 1.63, 2.96). This kind of effect modification finding (age modifies the effect of rural residence) would need to be checked for statistical significance (we can tell it would be, since the 95% CI of the two stratum-specific estimates do not overlap), and also to be sure the difference makes sense (preferably, it would have been hypothesized in advance). If we had not hypothesized this finding in advance, we’d be cautious about interpreting it as effect modification. Perhaps we’d recommend this for further follow-up in future studies. See Tables 8 and 9 for the stratum-specific results.

Table 8  Risk of disability by urban/rural residence. Younger adults (aged 18–64)

<table>
<thead>
<tr>
<th>Residence</th>
<th>Limited/disability?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural</td>
<td>Yes 132, No 840</td>
</tr>
<tr>
<td>Urban</td>
<td>Yes 314, No 1,244</td>
</tr>
<tr>
<td>OR = 0.62 (95% CI 0.50, 0.78)</td>
<td></td>
</tr>
</tbody>
</table>

Table 9  Older adults (aged 65+): risk of disability by urban/rural residence

<table>
<thead>
<tr>
<th>Residence</th>
<th>Limited/disability?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural</td>
<td>Yes 210, No 190</td>
</tr>
<tr>
<td>Urban</td>
<td>Yes 128, No 254</td>
</tr>
<tr>
<td>OR = 2.19 (95% CI 1.63, 2.96)</td>
<td></td>
</tr>
</tbody>
</table>
The results of the cross-sectional study above seem to suggest that rural residence is associated with an increased risk of disability in older adults, and protective in younger adults. What kind of study would make the causal inference stronger for this hypothesis? Since we cannot assign people randomly to their geographic residence, this will have to be an observational design.

**Answer:** These cross-sectional data raise concerns about causal inference because we don’t know if residence preceded disability. There may be reasons that that people may have moved to a different kind of setting. It is plausible that a person might move to a city for its healthcare, social, or transportation services if they had a disability. If so, the prevalence odds ratio we calculated above underestimates the association of rural living and disability. Or, perhaps, people with disability might stay in the rural area and be less likely to move because they had strong social support in the rural area. In this case, the lower prevalence of disability in urban areas might be falsely high and the rural excess inflated because of the exodus of people without disability. Whether or not either is true, they both pose potential problems to thinking that rural residence causes disability in older adults because of inconsistent temporality of the exposure (residence) and outcome. A stronger study design would be to examine cohorts of people, initially free of disability, to see which group was more likely to incur a disability. This is likely to be a question also that could benefit from a less heterogeneous outcome: for example, maybe sensory impairments are more likely to occur in one setting or another, but mobility impairments may be similar. Cohorts are, unfortunately, very expensive study designs; a very specific hypothesis about residence (the components of rural living that increase the risk of disability, for example) would probably take place over a long time period and require a large sample size and many different types of rural and urban settings.

**Relative Risk, Classification, and Confounding**

Andresen and Brownson (2000) analyzed the risk of disability among different ethnic groups of women in the United States. The data were based on a national stratified random-digit-dialed sample of women aged 40 and older; most questions were derived from the modules of the Behavioral Risk Factor Surveillance System, including disability definitions. One measure of disability was defined from a woman’s report that she was “limited” and also that she needed personal care assistance with activities such as eating, bathing, dressing, or getting around the house (classified as having activities of daily living dependence, or ADL). In another analysis, women were asked to describe their overall health status as excellent, very good, good, fair, or poor. Tables 10 and 11 show the responses of women who were Native American or Alaskan natives compared to white women. Participants included 774 white and 739 Native American women; because of some missing responses, not all analyses included all women who were interviewed. Are Native
American women more likely to be disabled according to ADLs? Are they more likely to be in fair/poor health (for this problem, answers are grouped as fair/poor vs. good-excellent)? Can you think of reasons that these results might differ? Consider that the average age of White women was 57.3, and that of Native American women was 54.4 (t-test for difference \( p < 0.01 \)). What additional analysis would you recommend to make sure these results were not biased by age? Why (Tables 10 and 11)?

**Table 10** Risk of ADL dependency for ethnic minority women

<table>
<thead>
<tr>
<th></th>
<th>ADL dependent?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Native women</td>
<td>25</td>
</tr>
<tr>
<td>White women</td>
<td>13</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>38</td>
</tr>
</tbody>
</table>

**Table 11** Risk of lower health status for ethnic minority women

<table>
<thead>
<tr>
<th></th>
<th>Fair/poor health status?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Native women</td>
<td>157</td>
</tr>
<tr>
<td>White women</td>
<td>156</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>313</td>
</tr>
</tbody>
</table>

**Answers:** In Table 10, the relative risk (RR) for ADL dependency for Native American women is 3.26 (95% confidence intervals [CI] of 1.69, 6.32; statistical significance, chi-square\(_{MH}\)<0.01) compared to White women. In Table 11, Native American women are more likely to be in fair/poor health, but the estimate is not as large (RR=1.70 and 95%CI 1.41, 2.06; statistical significance, chi-square\(_{MH}\)<0.01). Since these are prevalence data, some might argue for analysis using the odds ratio (OR); however, race/ethnic group is clearly temporally prior to the health status/disability determination, so we are less concerned about the problem of cross-sectional data here.

It is hard to compare directly the two RR estimates; while ADL dependency is one appropriate classification for disability, it is not synonymous with health status. Therefore we might expect that disability and health status have a relationship, but that they would not be overlapping definitions. A woman might require ADL assistance for mobility impairment, but consider she health status to be excellent or very good.

The problem of confounding by age is very likely here. Since various chronic conditions and mobility impairments increase with age (a prime example is arthritis), and Native American women are, on average, younger than White women in this sample, our calculated estimates of the risk of poor outcomes may be confounded by age. If we conducted a stratified analysis with age,
we would expect that the RRs we calculated may be biased low and that the age-adjusted RRs should be somewhat larger.

**Classification and Proxy Response**

In a reliability study of adults with disability and proxy respondents, we found that people with disability (PWD) and family members – who also answered for them – disagreed about the level of functional impairment and the pain experience of the PWD themselves (Andresen et al., 2001). Tables 12–14 show how their answers compare for one measure of dependence (needing help bathing) and the report of pain as a secondary condition. For each outcome variable, calculate the percent agreement, Kappa, and difference in the response (proportion) of the proxy from the person with disability. Why might the reliability be different for these two different variables?

<table>
<thead>
<tr>
<th>Table 12</th>
<th>Agreement of people and their family member proxies on dependence in bathing (needs any help)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PWD response</td>
</tr>
<tr>
<td>Yes</td>
<td>21</td>
</tr>
<tr>
<td>No</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 13</th>
<th>Agreement of people and their family member proxies on whether or not the PWD had pain during the prior year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PWD response</td>
</tr>
<tr>
<td>Yes</td>
<td>17</td>
</tr>
<tr>
<td>No</td>
<td>13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 14</th>
<th>Agreement of people and their family member proxies on whether or not the PWD had depression during the prior year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PWD response</td>
</tr>
<tr>
<td>Yes</td>
<td>12</td>
</tr>
<tr>
<td>No</td>
<td>13</td>
</tr>
</tbody>
</table>

**Answers:** The summary of these answers is listed below (and calculation of Kappa also described). Proxies do agree with PWD about needing assistance with bathing, although the Kappa is not in the “excellent” (above 0.75). Overall, their responses over estimated the need for assistance compared to the person with a disability. They also over estimated the depression of the PWD, and over estimated their pain. These results are common to tests of proxy response: the reliability is better for objective compared to subjective variables. The direction of differences also is common: proxies
consider the PWD to be more “disabled” then they are. However, some studies of proxies also suggest and underreporting of pain by proxies.

### Table 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>% agreement</th>
<th>Kappa</th>
<th>PWD % yes responses</th>
<th>Proxy % yes responses</th>
<th>Proxy difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need assistance bathing?</td>
<td>84.3</td>
<td>0.66</td>
<td>33.8</td>
<td>36.4</td>
<td>+2.6%</td>
</tr>
<tr>
<td>Pain last 12 months?</td>
<td>68.5</td>
<td>0.33</td>
<td>36.8</td>
<td>39.5</td>
<td>+2.7%</td>
</tr>
<tr>
<td>Depression last 12 months?</td>
<td>75.0</td>
<td>0.39</td>
<td>23.7</td>
<td>32.9</td>
<td>+9.2%</td>
</tr>
</tbody>
</table>

Calculating formulas for percent agreement and Kappa (for a completed treatment of these methods, see Fleiss, 1981).

(a) Calculating overall percent agreement

<table>
<thead>
<tr>
<th>PWD response</th>
<th>Proxy response</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>Total = 1.0</td>
</tr>
</tbody>
</table>

Percent agreement, or percent observed, is \( P_0 = P_{11} + P_{22} \).

But this is misleading because some agreement is due to chance alone. We therefore calculate the expected agreement (by multiplying the column and row totals for each cell, as in calculating chi-square statistics) and then summarize the agreement that is beyond chance, as a proportion of all that is possible. That defines the Kappa statistic (below). The examples are calculated below.

**Example 3a (Table 12): Agreement of People and their Family Member Proxies On Dependence in Bathing (Needs any help)**

<table>
<thead>
<tr>
<th>PWD response</th>
<th>Proxy response</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>1.0</td>
</tr>
</tbody>
</table>

Observed agreement is \( P_0 = 0.272 + 0.571 = 0.843 \).

Or, about 84% of the proxy-PWD sets agree on whether or not the PWD needs assistance in bathing.

**Example 3b (Table 13): Agreement of people and their family member proxies on whether or not the PWD had pain during the prior year**

<table>
<thead>
<tr>
<th>PWD response</th>
<th>Proxy response</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td></td>
<td>0.224</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
<td>0.461</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Observed agreement is $P_0 = 0.224 + 0.60 = 0.684$.

Or, about 68% of the proxy-PWD sets agree on whether or not the PWD had pain.

**Example 3c (Table 14):** Agreement of people and their family member proxies on whether or not the PWD had depression during the prior year.

<table>
<thead>
<tr>
<th>PWD response</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0.158</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>0.592</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Observed agreement is $P_0 = 0.158 + 0.592 = 0.750$.

Or, 75% of the proxy-PWD sets agree on whether or not the PWD had depression.

(b) Calculating expected agreement

<table>
<thead>
<tr>
<th>PWD response</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>$P_1$</td>
<td>$P_2$</td>
<td>$P_1$</td>
</tr>
<tr>
<td>No</td>
<td>$P_2$</td>
<td>$P_1$</td>
<td>$P_2$</td>
</tr>
<tr>
<td>Total</td>
<td>$P_1$</td>
<td>$P_2$</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Percent expected is $P_e = P_1^* P_2^* + P_2^* P_1^*$.

**Example 3b (Table 12):** Agreement of people and their family member proxies on dependence in bathing (needs any help)

<table>
<thead>
<tr>
<th>PWD response</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0.123</td>
<td>0.338</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>0.421</td>
<td>0.662</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.364</td>
<td>0.636</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Percent expected agreement is $0.123 + 0.421 = 0.544 = P_e$.

Or, over 50% of the agreement is expected by chance alone!

**Example 3b (Table 13):** Agreement of people and their family member proxies on whether or not the PWD had pain during the prior year.

<table>
<thead>
<tr>
<th>PWD response</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0.145</td>
<td>0.368</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>0.382</td>
<td>0.632</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.395</td>
<td>0.605</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Percent expected agreement is $0.145 + 0.382 = 0.527 = P_e$. 
Or, 53% of the agreement is expected by chance alone!

**Example 3c (Table 14):** Agreement of people and their family member proxies. On Whether or not the PWD had depression during the prior year.

<table>
<thead>
<tr>
<th>PWD response</th>
<th>Proxy response</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0.078</td>
<td>0.237</td>
</tr>
<tr>
<td>No</td>
<td>0.512</td>
<td>0.763</td>
</tr>
<tr>
<td>Total</td>
<td>0.329</td>
<td>0.671</td>
</tr>
</tbody>
</table>

Percent expected agreement is $0.078 + 0.512 = 0.590 = P_e$.

Or, 59% of the agreement is expected by chance alone!

(c) **Calculating Kappa:** Kappa is the measure of “beyond chance” agreement. That is, accounting for chance, how much more agreement do we observe (as a proportion of 100 better than chance)?

\[
\text{Kappa} = \frac{(P_o - P_e)}{(1 - P_e)}
\]

For our first example, agreement about dependence in bathing, this would be: \((0.843 - 0.544)/(1.0 - 0.544) = 0.66\). The Kappa is 0.66, or agreement of PWD and their proxies is 65% better than chance. This would be considered “good” agreement.

For our second example, agreement about pain, this would be: \((0.684 - 0.527)/(1 - 0.527) = 0.334\). The Kappa is 0.33, or agreement of PWD and their proxies is 33% better than chance. This would be considered “poor” agreement.

For our third example, agreement about depression, this would be: \((0.750 - 0.590)/(1 - 0.590) = 0.390\). The Kappa is 0.39, or agreement of PWD and their proxies is 39% better than chance. This would be considered “poor” agreement.

**Study Designs**

For each of the following research questions, consider the issues of (1) frequency of the outcomes, (2) feasibility and practicality, in measuring the outcome and/or exposures, (3) the issues of resulting causal inference, (4) the stage at which the question is directed (descriptive, hypothesis generating, hypothesis testing), and (5) the potential for sources of existing data. Suggest the best study design and explain the reasons for your choice.

(a) A new wheelchair allows people with mobility impairments to climb stairs, travel on soft terrain (e.g., sand), and physically rise to the head-level of five to six feet. It is expensive, but the developers think it will change the participation level, and quality of life of wheelchair users.

(b) Cleft lip (with or without cleft palate) occurs in about 1 to 2 births per 1,000 in Northern European countries and in people of these backgrounds in the United States. How would you investigate the hypothesis that a woman’s exposure to certain medications during early pregnancy may increase the risk of these birth outcomes?
(c) Cleft lip is usually resolved by surgery soon after birth or early in a child’s life. A relatively new surgery is recommended based on pilot work that shows children’s speech development is markedly improved compared to usual surgical techniques. The first studies occurred at only two academic centers, where candidates for surgery were highly selected. Speech development was followed only to the age of 30 months of age. Now the surgery has become relatively more common, and many more centers have elected to offer it to parents of children with cleft. The cost of the new procedure is high, and not all insurers will pay the additional costs. What kind of study would you recommend to help resolve the issue of improved speech outcomes?

(d) In planning for clinical and other services for children with asthma in your state, the Health Department and Educational Board need to know what the prevalence of severe asthma is (requiring ongoing clinical and/or hospital care) for children in the state between the ages of five (or first year of school) up to age 18 (or through high school). They will use these data to ask the state legislature to fund a new program of screening and case-management in the state’s schools. The legislative committee will meet in 3 months (they now are on their summer break). What kind of study would you conduct to meet the evaluation need and the time frame? The legislature is convinced there is a problem, but asks for a more thorough evaluation of “current need” in the State’s schools. They are willing to fund a study for up to $250,000 in the upcoming school year. How would you answer this question to help them determine need? What kinds of measurements/variables might you suggest for the study?

(e) Among adults with high-level spinal cord injury (SCI – affecting motor control of upper limbs), there is a large amount of variability in the incidence of upper respiratory infections (URI). While there is evidence that URI is increased compared to the general population, it is not clear if certain kinds of URI’s are more common or if URI is just more serious when it does occur (bringing it to medical attention). Because you are part of a large, pre-paid healthcare plan, you have access to a large clinical group of people with the appropriate SCI classification, and (1) can assume each person will attend a general medical clinic at least once a year, and (2) their other clinical, emergency, and hospital visits are available in the same medical care system; what kind of study would you perform to determine more exactly the nature and risk factors for URI in people with SCI compared to the other plan enrollees?

Answers

(a) This is a study that will provide some difficulty for “causal inference” in the most objective sense. The description of the new device is remarkable in its promise of providing more mobility in environments and situations where standard equipment fails. While a randomized trial of the device (new vs. standard wheelchair) is possible, and would provide evidence of changes and improve-
ments in participation, this may be a situation where one might use a staged random design: meaning a small number of people would receive the new wheelchair early (randomly), and the remainder of the sample would receive it later. This may help with ethical concerns and also allow fewer subjects to be enrolled. Outcomes might be compared (between groups) according to (1) levels (psychometrically tested scales) of participation, based on the ICIDH2; (2) changes in QALYs (quality-adjusted life-years), especially based on cost-per-QALY; (3) life satisfaction (based on psychometrically tested scales); and (4) traditional, function scales, like basic and Instrumental Activities of Daily Living (ADLs and IADLs), and the FIM, a commonly used measure of independence for rehabilitation research. Other issues that might be of interest include safety, social and or family issues, occupation/employment differences, and changes in repetitive motion injuries and pain.

(b) This fairly uncommon outcome may best be studied early on by a case control study. Medications taken during pregnancy may be recalled with some accuracy by mothers; in this design. Considerable care would need to be taken to assure that case mothers were not better at recalling their histories (or telescoping in exposures at other times); a validity study, perhaps using medical records and prescription records, would assist in finding out the overall accuracy of reports, and if it were “differential” by case status. Because pregnancy is of short duration, it also may be possible to do a cohort study; this would be especially true for exposures of interest that are uncommon (e.g., a specific medication used to treat infections). Any of these designs would be somewhat difficulty if asking about over-the-counter medications; however, diligent work on obtaining exposure information and/or getting women to record such information (specifically based on interviews at their first prenatal visit, for example), might overcome these difficulties.

(c) Since this elective surgery already is used, it may be difficult to mount a randomized trial. Parents, once knowledgeable, may naturally object to having the surgery withheld. However, if possible within a setting with no current payment for these surgeries (and the opportunity for the insurer and other sources to pay), this might be possible. The follow-up time that this requires (perhaps to school-age, a natural time to examine speech), would be prohibitive. For observational designs, a retrospective cohort would be a good solution. Within payer/insurance organizations, it would be possible to find all children born with cleft, the kind of surgery performed, and follow up to the present o perform in-depth language and speech testing at various attained ages of the children. Because of the strong potential link of SES (income) to the exposure, and the association this has on speech development, confounding of this and other factors would be a major task in such a study. The rareness of the exposure and difficulty in classifying cases (lower speech development) makes a case-control study impractical.

(d) If there were any sort of surveillance data (e.g., NCHS NHIS data) that can be used at your state, the fastest way to get at prevalence is to examine these data. Since 1998, the NHIS has provided sufficient data for each state to allow some state estimates. To provide these data, you would need the assistance
of a statistical analyst with strong skills in small area estimation and the NHIS. Without surveillance data (you could be lucky, and school absence records may give you some clue, for example), you would be very hard-pressed to provide these data. (2) The money might best be used to sample schools, then grades/classrooms, and then collect two kinds of data collection. First, a parent survey on asthma, asthma symptoms, diagnoses, healthcare visits among children (especially emergency room and hospital use), and school absences. Second, a test of asthma among students randomly selected would help address the issue of “undiagnosed/treated” disease. This would be expensive, requiring lung function tests in the field. As the potential staff for a program of screening and case-management, it would be prudent to fully involve school health personnel (e.g., nursing staff, parent volunteers, office staff).

(e) This might be a good opportunity to use the records and billing data of healthcare plan in a cohort study; potentially much of this work could be accomplished by database analysis, with no further data collection from individuals. If substantial data cleaning or database construction is required (e.g., combining pharmacy records, office visit records, hospital billing, etc), you might want to use data on all enrollees with SCI and a sample of others. An alternate, or auxiliary effort could be a survey of enrollees classified by exposure, as (1) all enrollees with SCI, compared with (2) a sample of other enrollees, matched possibly for age, gender. They would be surveyed, perhaps on several occasions, about their incidence of URI, symptoms severity, office and/or hospital visits. In either design, the potential for differential misclassification and identification of URIs exists. One would want to ascertain data accuracy (e.g., respirator infections noted and coded accurately, and the same for SCI and others?), and a validity sub study may be needed. The direction probably is in the direction of better ascertainment for enrollees with SCI, but you would want to confirm this.

**Classification Validity and Corrections to Prevalence Estimates**

In 1999, the Behavioral Risk Factor Surveillance System (BRFSS) had special questions on disability added to eight states and DC. From these data, it is possible to compare a narrow “gold standard” classification of mobility impairment to one your state has been using as an indirect estimate. In most states, the question “Are you limited in any way in any activities because of any impairment or health problem?” estimates disability prevalence. The most common reasons for affirmative responses are mobility-related (e.g., arthritis, orthopedic problems, low-back pain, etc). In the disability supplemental questions, you classify people as having “substantial” aid-assisted mobility impairment if they report they use any mobility aid (cane, walker, crutches, wheelchair or scooter, brace, artificial leg). Table 15 classifies 25,536 adult respondents.
(a) Calculate the sensitivity, specificity, and predictive value of a “yes” (positive) answer on the “limits” question vs. the narrow gold standard for mobility impairment.

(b) In your State, the prevalence of “limits” is 17.2%. Correct this estimate for the true prevalence of mobility impairment (with aids) for your State.

Calculation of sensitivity, specificity, and predictive value of a positive test (PV+) are via the formulas below:

\[
\begin{array}{c|cc|c}
\text{Gold standard (truth)} & \text{Yes} & \text{No} & \text{Totals} \\
\hline
\text{Yes} & a & b & a+b \\
\text{No} & c & d & c+d \\
\hline
a+c & b+d & \text{Total} \\
\end{array}
\]

\[
sensitivity = \frac{a}{a+c} \\
specificity = \frac{d}{b+d} \\
PV+ = \frac{a}{a+b}
\]

\[p + \text{specificity} - 1\]

\[
P = \text{sensitivity} + \text{specificity} - 1
\]

where: \(P\) = the true % of respondents with the characteristic; \(p\) = the measured % of respondents with the characteristic; Sensitivity and Specificity are the (known) parameters of the “p” measurement tool; in our case, we want to calculate \(P\) and we know that \(p\) in our state is 17.2%.

**Answers**

Answer a.  
Sensitivity \(= \frac{878}{1155} = 0.76(76\%)\)  
Specificity \(= \frac{21145}{24381} = 0.867(87\%)\)  
PV+ \(= \frac{878}{4114} = 213(21\%)\)
Answer b. \[
P_{\text{mobility impairment use}} = \frac{0.172 + 0.867 - 1}{0.76 + 0.867 - 1} = \frac{0.0390}{0.6270} = 0.0622(6.2\%)
\]

Acknowledgment Supported, in part, by funding from the Centers for Disease Control and Prevention (CDC; grant # U48/CCU710806) for the Methods Core of the Saint Louis University Prevention Research Center. The author gratefully acknowledges the primary mythological coursework that provides the foundation for this chapter from Drs. Noel Weiss and Thomas Koepsell at the University of Washington School of Public Health (Koepsell & Weiss, 2003). However, they bear no responsibility for errors or interpretations that deviate from their teachings. The author also appreciates the assistance of the following individuals in the preparation of this work: Tori Vahle, M.P.H., Janet Tang, M.P.H., Tricia McLendon, M.P.H., Erin DeFries Bouladin, M.P.H., and Tegan Boehmer, Ph.D. The students of my graduate course in disability epidemiology at the Saint Louis University School of Public Health also furnished valuable comments, questions, and editing; their time and patience have made this a better product.

References


Chan, E., Zhan, C., & Homer, C. J. (2002). Health care use and costs for children with attention-deficit/hyperactivity disorder: national estimates from the medical expenditure panel survey. Archives of Pediatric and Adolescent Medicine, 156, 504–511.


Minkler, M., Fadem, P., Perry, M., Blum, K., Moore, L., & Rogers, J. (2002). Ethical dilemmas in participatory action research: a case study from the disability community. *Health Education and Behavior, 29*, 14–29.


Public Health Perspectives on Disability
Epidemiology to Ethics and Beyond
Lollar, D.J.; Andresen, E.M. (Eds.)
2011, XXVII, 261 p., Hardcover