Research Design in End-of-Life Research: State of Science

Linda K. George, PhD

Purpose: The volume of research on end-of-life care, death, and dying has exploded during the past decade. This article reviews the conceptual and methodological adequacy of end-of-life research to date, focusing on limitations of research to date and ways of improving future research. Design and Methods: A systematic search was conducted to identify the base of end-of-life research. Approximately 400 empirical articles were identified and are the basis of this review. Results: Although much has been learned from research to date, limitations in the knowledge base are substantial. The most fundamental problems identified are conceptual and include failure to define dying; neglect of the distinctions among quality of life, quality of death, and quality of end-of-life care. Methodologically, the single greatest problem is the lack of longitudinal studies that cover more than the time period immediately before death. Implications: Gaps in the research base include insufficient attention to psychological and spiritual issues, the prevalence of psychiatric disorder and the effectiveness of the treatment of such disorders among dying persons, provider and health system variables, social and cultural diversity, and the effects of comorbidity on trajectories of dying. Key Words: Quality of dying, Death, Dying

Research on dying persons, their families, and end-of-life care has grown dramatically in volume during the past quarter century. There are undoubtedly multiple reasons for this increase. Among them are the shift in priorities and concerns resulting from a population that is aging rapidly in both the number and proportion of older adults, policy concerns about escalating health care expenditures, especially in late life, and public perceptions that dying too frequently lacks adequate control of pain and suffering, as well as basic human dignity.

The potential for empirically grounded knowledge to inform these issues depends on both the quantity and quality of the research available. The sheer volume of research is important, but, with public policies and treatment guidelines at issue, the research also must be rigorous and generalizable. The purpose of this article is to review the quality of end-of-life research to date. This article is divided into three sections: (a) a brief description of the methods used in sampling the research base, (b) an assessment of the methodological adequacy of end-of-life research to date, and (c) a broader discussion of opportunities to improve the quality of end-of-life research, as well as high-priority topics for future research.

Methods of This Review

Although it is only recently that end-of-life research has received substantial attention, the relevant research base is broad. Publications included in this review were identified by two methods. The first was a key word search of primary bibliographic databases in the clinical, behavioral, and social sciences. To be appropriately inclusive, multiple key words were used, including “end-of-life,” “terminal illness,” “death and dying,” “palliative care,” and “psycho-oncology.” The three primary databases used were Medline, PsychInfo, and the Social Science Citation Index. These sources generated a pool of more than 1,000 articles. Although date of publication was not restricted, approximately 90% of the articles were published after 1990, documenting the increasing scientific attention to death, dying, and end-of-life care. In addition, several experts in the field provided their bibliographies of end-of-life research.

To be included in this review, an article had to include empirical findings. With this constraint, the sample of articles used in developing this article was reduced to approximately 400. The author reviewed all of these articles. The major types of studies that emerged from this selection process include randomized clinical trials; uncontrolled trials; surveys of patients, family members, and health care providers on a variety of end-of-life issues; epidemiological studies,
Conceptual Quandaries

Two other exclusion criteria were used. First, the review was restricted to articles found in journals printed in English. Most of the articles used American samples, but there were also significant numbers of studies from Britain, Canada, and Australia and a few studies from Scandinavian countries and Israel. Second, studies of educational interventions that were evaluated solely by means of tests of knowledge were excluded. However, studies of educational interventions that included measures of change in behavior (e.g., provider communication, willingness of patients to put advance directives in place) are included.

State of Science in End-of-Life Research

Although much end-of-life research is of excellent quality, highlighting those studies is not the focus of this review. Instead, the primary focus is on issues that have been inadequately handled or ignored completely in research to date. One should not view these comments, however, as indictments against all studies. This section addresses three issues: conceptual concerns, basic research design, and sampling.

Conceptual Quandaries

Conceptual clarity is a limiting factor in all research. No study can be better than the conceptual framework used to anchor it; every aspect of research design and implementation is inextricably affected by the conceptual foundation that underlies the research question. Although well-articulated conceptual frameworks guide many end-of-life research efforts, three conceptual issues seem to be pervasive and highly limiting with regard to the quality and scope of empirical work to date.

Defining Dying.—The single most difficult and poorly handled issue in end-of-life research is the mismatch between conceptual and operational definitions of “dying” or “terminal illness.” A wide variety of operational definitions, often implicit rather than explicit, are found in research to date. For some investigators, individuals are categorized as dying or terminally ill on the basis of diagnosis alone. Other investigators use prognostic criteria consisting of varying combinations of diagnosis, symptom expression, and functional capacity to categorize individuals as terminally ill. Most frequently, investigators rely on physicians’ prognoses about life expectancy, and this is especially true in randomized controlled trials, despite empirical evidence that physician prognosis is typically highly inaccurate (e.g., Christakis & Lamont, 2000; Fox et al., 1999). But I also found studies that relied on the prognostic assessments of the patients or their family members to define whether the patient was dying. In some studies, the health care setting is used to define who is dying—a significant number of studies use receipt of hospice care, palliative care, or even being in the intensive care unit (ICU) as a definition of “dying.” Also, in a large proportion of studies, no definition of “dying” or “terminal illness” is offered or can be discerned implicitly.

In contrast to these operational definitions of “dying” are a variety of conceptual models of the dying process. In a recent excellent review, Copp (1998) identified six categories of theoretical frameworks: (a) stage theories of dying (e.g., Buckman, 1993; Kubler-Ross, 1969); (b) Glaser and Strauss’ (1965) context of awareness theory; (c) dying trajectory theories, which range from the classic “time for dying” theory of Glaser and Strauss (1968) to the superb analysis of Medicare decedents that Lunney, Lynn, and Hogan (in press), prepared for this conference; (d) Pattison’s (1977) living-dying interval theory of dying; (e) task-based theories of dying (e.g., Corr, 1992; Kalish, 1979); and (f) Copp’s (1997) readiness-to-die theory. Space limitations preclude discussion of the details, similarities, and differences among these types of theories. The important points for the purpose at hand are that (a) there exist a variety of richly woven theories of the dying process, and (b) these theories have received little empirical testing other than the testing that generated the theory. Certainly, these theories are seldom acknowledged and rarely, if ever, inform the myriad of end-of-life studies that have appeared in the past decade.

There are important implications of the failure of the field to develop congruent conceptual and operational definitions (or set of definitions) of “dying” or “terminally ill.” First, the absence of theoretical foundations for empirical research precludes the development and testing of specific hypotheses; moreover, most end-of-life research includes no hypotheses. Second, from a methodological perspective, if the operational definition of the relevant population is ignored or “fuzzy,” it is impossible to evaluate other components of the study, such as sampling. Consider, for example, what Teno and Coppola (1999) call the “problem of the denominator.” Unless definitions of who is and is not dying are conceptually and operationally clear, the population cannot be defined. Without a clearly defined population, important kinds of conclusions are precluded, even if the study is rigorous in every other way. Third, with the profusion of operational definitions and nondefinitions of this critical concept, it is impossible to compare findings across studies with accuracy or confidence. Finally, in the absence of clear conceptual definitions, accompanied by rigorous measurements based on those definitions, it is difficult to see how the research base can be confidently used to inform practice or policy.

In short, I view the absence of conceptual and operational congruity regarding definitions of “dying” and/or “terminally ill” as the most important issue facing end-of-life research. I cannot see the field breaking new ground or “reaching the next level” without resolving this issue.

Distinguishing Between Quality of Life and Quality of Dying.—Another conceptual quandary in end-of-life research concerns the distinction between quality
of life and quality of dying. These are not synonymous terms, although they are often used as such, either explicitly or implicitly. Conceptually, it is simple to see that these are different phenomena. One can easily imagine two terminally ill patients, both of whom report low quality of life, but one of whom rates his or her quality of dying as high, whereas the other rates it as low. It is possible, of course, that measures of quality of life and quality of dying would be highly correlated. However, (a) we cannot empirically address that issue until there are conceptually distinct definitions and operationally distinct measures of them, and (b) evidence of a high correlation between them would not mean that they are the same thing. (For example, there is a high correlation between cancer and chemotherapy, but they are certainly not the same thing.)

At the conceptual level, the concept of quality of life has been transferred, largely intact, from other areas of clinical and behavioral research. There have been efforts to assess the psychometric properties of extant quality-of-life measures with samples of terminally ill patients (e.g., Hardy, Edmonds, Turner, Rees, & A’Herne, 1999; Hickman, Tilden, & Tolle, 2001), but that is different from reconceptualizing the concept. Several instruments also have been developed to specifically measure quality of life among terminally ill individuals (e.g., Byock, 1995; Cohen, Mount, Strobel, & Bui, 1995; McMillan, 1996). The latter instruments are generally superior to other measures of quality of life in that they include content domains that are typically not included in more general instruments, but are especially relevant to the terminally ill (e.g., spirituality, sense of meaning).

Efforts to conceptualize quality of dying are much fewer in number and, consequently, less sophisticated than those for quality of life. A small number of studies have attempted to conceptualize quality of dying, using the preferences and priorities of patients as the grist for the mill, and there is surprising consistency among patient reports in these studies. In the best study to date, Steinhauser and colleagues (Steinhauser, Christakis, et al., 2000; Steinhauser, Clipp, et al., 2000) interviewed terminally ill patients, their families, and a range of health providers about their views of a “good death.” Six primary domains were identified and recommended as salient dimensions of quality of dying: symptom and pain management, clarity of decision making, preparation for death, completion, contributing to others, and affirmation of the whole person. These domains are substantially different from those typically included in quality-of-life instruments and are congruent with issues raised by terminally ill patients and their families in qualitative studies (e.g., Barnard, Towers, Boston, Lambrinidou, 2000; Byock, 1997). Steinhauser, Clipp, and Tulsky (2002) also point out that the domains that emerged from their interview data tend to integrate topics that often are viewed as distinct dimensions in quality-of-life instruments. For example, completion may include financial arrangements, having intimate discussions with loved ones, and coming to accept one’s life as meaningful and good. The development and validation of quality of dying instruments are high-priority issues for future research. In the interim, it is important that quality-of-life measures not be interpreted as if they are synonymous with quality of death.

A related construct is quality of care for dying patients. Not all issues relevant to quality of dying can be directly addressed by health care professionals (although health care settings vary in the extent to which they are conducive to patients working with and through nonmedical concerns). On the other hand, it is important that end-of-life care address all of its relevant components as effectively as possible. Efforts to develop a Toolkit of Instruments to Measure End-of-Life Care are under way (Teno & Landrum, 1996). Teno, Byock, and Field (1999) argue that the instruments included in the Toolkit should be patient-focused, family-centered, clinically meaningful, administratively manageable, and psychometrically sound. This list of criteria appears to be highly appropriate, paying attention to the experiences of patients and families, high-priority clinical issues, and conventional standards of psychometric adequacy. Until well-documented instruments are available, it will be difficult to develop compelling comparisons across health care settings (e.g., hospice vs. hospital-based palliative care).

Depending on the research questions under investigation, one, some, or all of these measures may be relevant. As seriously ill patients are followed over time, both quality-of-life and quality-of-dying instruments are likely to be necessary to monitor changing concerns and capacities. One would hypothesize that quality of dying would be associated with quality of care at the end of life; thus, it would be important to measure both of them in some studies. It is most important that these constructs be kept conceptually and operationally distinct. Further investment in both developing valid measures of quality of dying and quality of end-of-life care and validating existing instruments with large samples will be required before the field can proceed in these directions.

**Overreliance on Consumer Satisfaction.** —An impressive research base focuses on quality of life, psychological well-being, and other subjective, experiential states of aging individuals. This research compellingly documents the importance of including subjective perceptions as outcome measures for evaluating both experimental interventions and the consequences of naturally occurring phenomena, such as acute and chronic stressors and aging itself. Nonetheless, subjective perceptions should seldom be the only outcome examined. Similarly, although consumer satisfaction is unquestionably an important component in evaluating quality of care at the end of life, it should not be the only criterion by which such care is measured. In end-of-life research to date, consumer satisfaction—sometimes patient satisfaction, but more frequently family satisfaction—was often the only outcome measure examined. Overreliance on measures of consumer satisfaction...
satisfaction has several significant limitations and even dangers.

First, and at the most global level, it is well known that satisfaction scores are not a simple reflection of levels of objective well-being or conditions. This has been repeatedly demonstrated in a variety of areas. For example, most older adults overestimate their physical health, compared with objective assessments (e.g., Borowski, Kinney, & Kahana, 1996). Similarly, approximately half of the individuals age 65 and older who have incomes at or below the poverty level report that they are satisfied with their incomes and standard of living, whereas 16% of Americans who earn more than $250,000 per year describe their incomes as inadequate for meeting their needs (George, 1992). These discrepancies between objective conditions and subjective perceptions of them are interesting and, for some scientific purposes, important. They certainly suggest that subjective perceptions alone should not be the basis for policy and practice—if it were, a large number of objectively poor people would be exempt from need-based programs, and our society would make regular income transfers to some of its wealthiest citizens. Less dramatically, but also relevant, subjective assessments are significantly related to personality traits and other characteristics (Okun & George, 1984) that I would argue should not be the basis for policy and practice.

Second, and as noted previously, some issues relevant to the quality of dying are not in the scope of medical treatment. Terminally ill individuals who are receiving state-of-the-art, end-of-life care may be troubled by existential despair, unresolved conflicts with or estrangements from family members, or guilt about previous behavior. Such persons are likely to express dissatisfaction with their quality of dying; yet, this dissatisfaction is unrelated to their medical care.

Finally, empirical evidence—limited in scope, but with fascinating implications—suggests that there are conditions under which ethical treatment will lower satisfaction. Research on end-of-life decision making, especially on the willingness of seriously ill persons to put advance directives in place, provides a useful illustration. From an ethical perspective, advance directives are important for at least two reasons: (a) they respect the dignity of the individual to make important personal decisions, and (b) discussion of advance directives is intended to educate individuals about their rights of self-determination. There is substantial evidence, however, that approximately half of seriously ill patients do not want the responsibility for end-of-life decisions—a pattern observed in both studies that survey patients (e.g., Hofmann et al., 1997; Hopp, 2000) and in clinical trials intended to educate patients about advance directives (e.g., Molloy et al., 2000). Most intervention studies report negative findings—that intervention does not increase advance directives. In one intervention study conducted to increase the use of advance directives among persons with HIV/AIDS, however, the investigators increased the use of advance directives, but that was accompanied by a decrease in patient satisfaction with health care (Ho, Thiel, Rubin, & Singer, 2000). In none of the studies available did increased knowledge about advance directives and other end-of-life options increase patient satisfaction (e.g., Molloy et al., 2000; Puchalski et al., 2000).

These findings raise questions about the outcome of interest. All of the intervention studies increased patients’ knowledge about options for self-determination in end-of-life care, but none of them increased patient satisfaction—and one lowered it. If consumer satisfaction was the sole criterion for evaluating end-of-life care, results to date would suggest that efforts to educate patients about their rights for self-determination and decision making should be eliminated. If, however, we view patient education that fosters self-determination as an appropriate goal of end-of-life care, we would probably reach the opposite conclusion.

A range of outcomes is relevant to understanding the experience of dying and the effects of end-of-life care (Devery, Lennie, & Cooney, 1999). Consumer satisfaction is an important outcome, but it should not be the only outcome of interest. Patient-centered care means care that not only places the desires of the individual at the top of the priority list, but also includes technically effective care and care that respects the dignity of the patient, whether or not that patient is satisfied. A range of outcomes should be used in future clinical trials and naturalistic studies.

Basic Research Design

For our purposes, basic research design refers to the overall design of end-of-life studies. Pertinent issues include (a) cross-sectional versus longitudinal designs, (b) patient surrogates, (c) the mode of data collection, and (d) the use of retrospective data.

The Need for Longitudinal Studies.—A major limitation of end-of-life research to date is the relative neglect of longitudinal designs. Certainly many studies, especially randomized clinical trials and other intervention studies, use longitudinal data. Unfortunately, however, even in these studies, the period of observation is typically short, and in some studies, only the outcome of interest is monitored over time (e.g., the decision to enter hospice, whether or not patients put advance directives in place).

As noted previously, virtually all theories of dying are process theories, although the hypothesized processes differ across theories. To date, none of these theories has had adequate testing using the longitudinal data required to test them. (One possible exception is Kubler-Ross’ [1969] stage theory, which has been largely dismissed, although it appears that logical criteria were more important than empirical findings in disputing this theory [e.g., Kastenbaum, 1985]). Even the relatively large number of studies that examine quality of life during terminal illness and consumer satisfaction with end-of-life care are overwhelmingly cross-sectional studies. As a consequence, we do not know how quality of life and
consumer satisfaction change (or not) over time and in response to changes in clinical status and/or treatment decisions.

One of the greatest needs in end-of-life research is increased attention to the process or trajectories of dying. In a ground-breaking article, Lunney and colleagues (2002) used Medicare claims data to develop four common trajectories of dying and were able to document distinctive differences across trajectories in demographic characteristics, diagnosis, patterns of care, and health care expenditures. Hopefully, this article will generate more studies of trajectories of dying that extend to topics such as quality of life, decision-making “turn points,” psychological distress, and spiritual concerns, as well as specific parameters of clinical care such as pain management. Very basic information about distinct pathways to death also will play a critical role in the development of future interventions to improve the quality of care of dying patients. Current intervention trials, most of which have reported disappointing results, have not been grounded in a multidimensional understanding of the natural history of terminal illness.

To be sure, significant obstacles confront investigators who wish to perform longitudinal studies of dying persons. For example, as Lunney and colleagues (2002) point out, one trajectory of dying is unanticipated, “sudden death.” Because these deaths occur without warning, only retrospective studies of these decedents will be possible. Even when individuals experience a trajectory of death that involves an identifiable decline, there are practical problems in conducting prospective studies. A major issue of concern is attrition, of which there are two kinds: attrition from mortality and what might be called “functional attrition.” Extant prospective studies of dying patients report high rates of mortality among their samples (e.g., Albert, Murphy, Del Bene, & Rowland, 1999; Hamel, Davis, et al., 1999). In part, of course, this is inevitable. Attrition problems also result in part from research design. A large proportion of the samples in end-of-life studies are recruited from hospices, palliative care units, and other health care settings that treat critically ill patients (e.g., ICUs). There is substantial evidence that patients typically do not enter these settings until death is imminent. Using data from a national sample of Medicare beneficiaries, for example, Christakis and Escarce (1996) found that the median survival of patients entering hospice is 36 days—a very short period of time for longitudinal studies of hospice patients. Although hospice, palliative care units, and similar settings offer relatively easy ways to identify samples of terminally ill patients, use of these recruitment sites ensures short-term survival of many of the patients (e.g., Christakis, 1994; Jordhoy et al., 1999; Travis, Loving, McClanahan, & Bernard, 2001).

“Functional attrition” refers to the fact that dying patients often are unable to participate in interviews and other forms of data collection for a period of time before death. Some investigators report that response rates fall precipitously within a week of death (e.g., Hinton, 1996a, 1999); others report that patients are often unable to participate in data collection for as much as a month before death (e.g., Jordhoy et al., 1999; Kreling, Wu, & Lynn, 2000). Again, recruitment from terminal care settings, such as hospice and palliative care units, ensures a high rate of functional attrition, in addition to high rates of mortality. An obvious priority for future research, especially prospective studies of terminal illness, is to develop alternate strategies for identifying and recruiting study participants. Alternate sampling frames will undoubtedly be logistically more difficult and more expensive, but failure to pursue such strategies results in a very incomplete understanding of dying.

Longitudinal studies that identify more long-term trajectories of dying also pose measurement challenges because instruments must cover the full range appropriate for studies of terminal illness. In general, instruments useful for longitudinal end-of-life studies must be equally robust in making meaningful and valid distinctions among (a) relatively functional persons who are in the early stages of illness or whose illnesses are responding to treatment; (b) persons in the middle stage of illness, in which prognosis is perhaps uncertain, and treatment patterns are complex combinations aimed at both symptomatic relief and curative options; and (c) persons in late-stage illness, probably suffering increased symptoms and decreased functioning, and facing impending death.

Use of Surrogate or Proxy Respondents.— Compared with other areas of clinical and behavioral research, use of surrogate or proxy respondents is very prevalent in end-of-life research. In part, use of surrogate informants reflects the fact that many dying patients are unable to participate in primary modes of data collection (e.g., interviews). Moreover, there is general consensus that end-of-life care must focus on the families of the terminally ill, as well as on the dying patients. Dying patients typically desire that their families play a dominant role in end-of-life care decisions (Hopp, 2000; Puchalski et al., 2000). For all these reasons, inclusion of family members in end-of-life research is appropriate.

Nonetheless, in studies to date (a) criteria for collecting data from patients versus surrogates are seldom articulated, and (b) surrogates are sometimes used in situations in which the patient is capable of participating in data collection. Almost totally missing are clear explanations of the rationales for choice of respondents and descriptions of the circumstances under which surrogates are used in lieu of dying patients. Moreover, it appears that surrogates are sometimes used to report on behalf of patients who are capable of reporting their own experiences. As noted previously, a relatively large body of research addresses satisfaction with end-of-life care. It is interesting to note that most of this research is based on the satisfaction of patients’ family members, rather than on the evaluations of the patients themselves. Clearly, few patients will be able to provide satisfaction ratings at the end of the dying trajectory, but investigators...
apparently feel no responsibility to explain why patient satisfaction is not ascertained when possible.

If patients and surrogates provided concordant information about the illness experience of the patient and their reactions to end-of-life care, there would be no reason for concern. There is sufficient research, however, to suggest significant mismatches between the reports of patients about their symptoms, levels of pain, and similar issues and those reported by their closest family members (e.g., Higginson & McCarthy, 1993; Hinton, 1996b; Layde et al., 1995). In addition, some research suggests that there is systematic bias in surrogate reports such that they report lower quality of life and greater discomfort than the patients for whom they are reporting (Higginson & McCarthy, 1993; Magaziner, Simonsick, Kasher, & Hebel, 1988). Comparisons of patient and health provider reports offer an interesting contrast. Nurses’ ratings of patient symptoms, pain, and distress are typically more congruent with those of patients than are family members’ ratings (Higginson & McCarthy, 1993; Nekolaichuk et al., 1997). In contrast, physicians are the least accurate surrogates and consistently underestimate patients’ symptoms and distress (Nekolaichuk et al., 1997; Stephens, Hopwood, Girling, & Machin, 1999).

Even when surrogate respondents are necessary or appropriate, criteria for selection of an informant are seldom reported. Many studies claim to have obtained data from the “primary caregiver” or “family member most involved with care,” but the criteria used to identify those family members are not described, and undoubtedly vary across studies.

These problems suggest three recommendations for future research in which surrogates are considered as informants for patients. First, if the research question concerns the experience of patients, investigators should make every reasonable effort to obtain information directly from the patient. Second, the information obtained from surrogate informants should be restricted to observable behaviors (e.g., patient requests for additional pain medication rather than ratings of the level of the patient’s pain). Some kinds of information, such as ratings of patient quality of life, cannot be accurately obtained from proxy respondents. Third, research protocols should include well-articulated criteria for the conditions under which information about the patient is obtained from surrogates, as well as for surrogate selection, and those criteria should be included in published descriptions of the research design.

**Mode of Data Collection.**—A variety of data collection strategies have been used in end-of-life research, including interviews, self-administered surveys, chart reviews, extracted information from death certificates and insurance claims, and direct observation of dying patients and/or end-of-life care. All of these modes of data collection are appropriate, given that the data collection strategy is dictated by the research question of interest. However, available research suggests possible overreliance on patients’ medical charts as sources of information.

The temptation to use medical charts as data sources whenever possible is understandable. Logistically, this is perhaps the easiest form of data collection. Medical charts are typically easily attainable, information may be recorded in a way that is useful for research purposes, and there are no time pressures for data collection. Nonetheless, medical charts are also notoriously incomplete and often inaccurate (e.g., Luck, Peabody, Dresselhaus, Lee, & Glassman, 2000). In addition to health care providers failing to record all relevant observations and treatments, reliance on medical charts will exclude the various contributions of informal providers. Family caregivers and other relatives often contribute to the care of dying patients in important ways, ranging from help with grooming, efforts to relieve pain (e.g., by massage, by changing the patient’s position), and efforts to reduce distress (e.g., Emanuel et al., 1999). This is true in hospitals, hospice settings, and nursing homes, as well as in instances in which the dying patient remains at home. Using medical charts as the primary data source for measuring end-of-life care also ignores both (a) the kinds of care and assistance that the patient received before admission and (b) care patterns experienced by patients who die at home.

**Use of Retrospective Data.**—Relative to other areas of clinical and behavioral research, end-of-life research appears to disproportionately rely on retrospective data, typically obtained from bereaved family members. A large proportion of the studies of consumer satisfaction with end-of-life care, for example, rely on family members’ retrospective reports. Of even greater methodological concern, many studies ask bereaved family members to report on issues such as the dying patient’s symptoms, level of pain, and psychological distress just before death (e.g., Addington-Hall & McCarthy, 1993; Hanson, Danis, & Garrett, 1997; Lynn et al., 1997). As noted previously, it is well documented that even when patient and surrogate reports are made at the same time, mismatches are common. An important, but limited, research base suggests that surrogate reports are even less accurate weeks or months after the patient dies than they were when collected while the patient was alive (Hinton, 1996b).

There are legitimate reasons to ask bereaved family members about their experiences with end-of-life care. End-of-life programs can use such information to determine whether there are issues of concern for significant proportions of families, and use that information to develop alternative care patterns. It also can be helpful to obtain family members’ perceptions after they have had time to gain perspective. This is substantially different, however, from eliciting postmortem information from family members and using it to make generalizations about the process of dying or other experiential facets of end-of-life care. In general, after a significant amount of time has passed, surrogate reports should be viewed as their current perceptions, not as an accurate report of the time of the family member’s death. The general perils of relying on...
retrospective data are well-known (e.g., Dex, 1995; Scott & Alwin, 1998). They need to be taken as seriously in end-of-life research as in other research domains.

Sampling

Sampling is a critically important element of research design. High-quality sampling strategies involve some kind of representative selection among the elements of an appropriately defined population (e.g., Kish, 1965). The overall quality of sampling in end-of-life research is technically very poor; research based on national samples (typically place-of-death studies and some studies of patterns of care and/or end-of-life health care expenditures) are important exceptions to this conclusion. There are several ways in which the sampling strategies used in most end-of-life research are deficient from the perspective of sampling theory.

First, and most importantly, the population of interest is seldom defined, let alone sampled adequately. End-of-life research suggests multiple populations of interest, including dying persons, dying persons receiving specific types of care (e.g., hospice care, palliative care, nursing home care, home care), persons dying of specific diseases, and persons who have and have not made specific end-of-life decisions (e.g., advance directives). In most studies, the population of interest is only implied, not articulated. If the population was articulated, it would highlight the fact that most investigators fail to sample systematically (let alone representatively) from the population of interest. Most studies of end-of-life care modalities, such as hospice or nursing home care, are based on patients using a single health care setting. There is no attempt to sample from the universe of such health care settings. Most studies of deaths resulting from specific diseases are persons from a given health care setting who died of the disease(s) of interest.

Even if investigators cannot systematically sample from the population of interest, it is important that they define that population. This informs readers of the broader population of interest and also provides a basis for considering the appropriateness of the sample for the research question. For example, if the research question concerns deaths resulting from stroke, use of a single type of health care setting (such as hospitals or nursing homes) will inevitably lead to biased results. Stroke patients are found in a variety of health care settings (including home), and there is every reason to expect that selection factors strongly influence location of care.

Second, as noted previously, investigators fail to define conceptually what they mean by “dying patients” or “terminally ill individuals.” In the absence of a conceptually clear definition linked to appropriate methods of case identification, the populations relevant to end-of-life research cannot be identified. What is the sense of defining the population of interest as “terminally ill nursing home patients” in the absence of a definition of what it means—conceptually and operationally—to be dying? If an investigator wishes to study dying nursing home patients, he or she must be able to reliably and validly categorize nursing home patients into those who are and are not dying—and then sample among the former.

A final issue of concern, from the perspective of sampling theory, is the process by which elements of the population are sampled. Random sampling, which can take several forms, is best because it protects against bias. I find no evidence of attempts to randomly sample populations of interest in most end-of-life research. Instead, investigators appear to select a place for subject recruitment (e.g., hospice settings, palliative care units) and then simply enroll as many patients as are willing to participate.

Sampling theory is the ideal, but it must be balanced against practical realities. It is quite clear that the ideal methods for defining a population, identifying the elements of that population, and then randomly sampling those elements are unrealistic for most end-of-life research. There are, however, important ways in which the quality of samples used in end-of-life research can be substantially improved and/or their limitations understood.

First, if possible, patients should be recruited from multiple sites, to permit generalizability beyond a single site. If an investigator is interested in trajectories of dying in nursing homes, multiple nursing homes should be included in the sample. If an investigator is interested in the terminal trajectories of malignant melanoma patients, patients should be recruited from multiple health care settings. Second, inclusion and exclusion criteria for sample members should be developed and applied. I seldom see this issue addressed in end-of-life research. If the sample is restricted to English-speaking persons, say so. If it is restricted to patients for whom there is a family member available, say so. Third, investigators should always (but often do not) report the participation rate achieved and, to the extent possible, compare the clinical and demographic characteristics of participants and non-participants. The participants in any given end-of-life study are a small “slice” of the population and an incomplete set of the potential participants. We need to know the characteristics of persons who will not participate in end-of-life studies, as well as those of participants.

This discussion of sampling has been described in the context of sampling dying individuals. The same issues apply to other relevant groups of study participants, including family members and health care providers. It also may be tempting to assume that these concerns apply only or primarily to in-person data collection. They also apply, however, to studies based on medical charts. Investigators using chart data also need to define their population, define the criteria used to select or reject charts, and the extent to which the data were available (e.g., Were there missing charts? How many charts had to be rejected because of incomplete information?). Finally, it must be acknowledged that some research traditions reject rigorous attention to sampling theory. This is especially true of investigators using qualitative methods. Although I would agree that random sampling is unlikely
in small, qualitative studies, the investigators of such studies should describe fully the process by which potential subjects were identified and recruited.

**Broader Recommendations**

This final section of this review presents a set of more general recommendations for future research. These recommendations are viewed as potential opportunities to advance our knowledge of the end of life and to simultaneously improve the methodological quality of the research base and areas of research that have been neglected or understudied.

**Opportunities**

**“Piggyback” on Existing Studies.**—One of the most pressing needs in end-of-life research is studies that include substantial amounts of information about what individuals are like before terminal illness and death. Most end-of-life research is based on samples whose members have been labeled as dying when data collection begins. This sampling strategy can generate useful information, but we also need studies that follow individuals over long periods of time before they face fatal illnesses or conditions. We currently know very little, for example, about how people’s attitudes about advance directives and other preparations for death change (or remain stable) as they move from relatively healthy states to the experience of dying. We know little about how relationships with loved ones change as both the dying persons and their family members recognize and accept impending death. Very importantly, if preterminal care data were available, we would be able to study the selection factors associated with treatment choices and treatment settings. These are illustrations of a large number of significant research issues that could be addressed if longitudinal data covering both the preterminal and terminal phases of life were available.

Important opportunities are available in this arena. A number of high-quality longitudinal studies of older adults are available or in progress. The National Institute on Aging (NIA), for example, funded four EPESE (Established Populations for Epidemiologic Studies of the Elderly) studies, each of which collected nearly a decade of extensive data on social factors, physical and mental health, cognitive status, health behaviors, and health care. NIA is now funding two major longitudinal studies of the aging population: the Health and Retirement Study and the National Long-Term Care Survey. It would be relatively inexpensive to augment those ongoing studies to (a) ask questions about end-of-life preferences and preparations; (b) obtain a set of standard data about the timing, causes, and circumstances of death; and (c) if desired, conduct brief interviews with the decedents’ family members. The cost of augmenting these studies would be minor in relation to the knowledge to be gained from having detailed information about participants before the terminal phase of life.

**Parallel Studies in Multiple Settings.**—Another pressing need in end-of-life research is to compare trajectories of dying across multiple settings. Currently, the only way to gain insight into the differing experiences of dying at home, in a nursing home, in a hospital, and in hospice is to compare results of studies in which the methods used were not comparable—and, even worse, in which investigators were not addressing the same research questions. (See Bruera, Neumann, Brenneis, & Quan, 2000, however, for a useful comparison of predictors of patient admissions to palliative care units, acute hospital wards, and hospices.) If we are to obtain rigorous data about the similarities and differences of dying as experienced in different places, studies are needed in which the same methods are used in multiple settings. Studies that compare trajectories of dying across multiple settings would yield methodological advantages as well, providing, for example, opportunities to establish the validity of measurement tools across distinct settings and samples.

**Learn From Other Research Fields.**—Another opportunity in end-of-life research is to make better use of challenges that have been confronted and resolved to some level of scientific adequacy in other fields. One of the challenges in end-of-life research is integrating and making sense of information obtained from multiple informants. Many studies obtain data from both dying patients and their surrogates; some studies obtain data from health care providers as well. Investigators who collect data from multiple sources face the difficult task of integrating the responses of multiple informants into meaningful analyses.

End-of-life researchers are not the first to confront the challenges of integrating data from multiple sources. Scientists who study children have faced the same issue for decades. In their studies, collecting data from both parents and children, about the same topics, is nearly universal, and in many studies, data are obtained from other sources as well, such as teachers or peers. Developmental psychologists have developed a series of strategies for integrating data from multiple informants, much of it incongruent, into meaningful analyses (e.g., Ablow et al., 1999; Mesman & Koot, 2000; Webster-Stratton & Hamond, 1998). A concerted effort to learn those strategies and assess their utility for end-of-life research would be worth the effort. There are undoubtedly other methodological issues that could be illuminated by examining the research of investigators who have confronted similar issues, albeit in different contexts.

**Gaps in the Knowledge Base**

**Underrepresentation of Psychological and Spiritual Issues.**—Compared with the volume of research examining pain, fatigue, and other physical symptoms during terminal illness, much less attention has been paid to psychological and spiritual issues—a pattern also observed by others (e.g., Francke, 2000; Johnston & Abraham, 1995). Only a few studies address
psychological distress and related issues during the last weeks or months of life. In a significant number of these, investigators relied on the reports of surrogates rather than patients. Moreover, very few studies were based on U.S. samples. Two of the most comprehensive U.S. studies, HELP and SUPPORT, included brief assessments of depression and related symptoms, but these data have received much less attention about physical symptoms (see, however, Somogyi-Zalud, Zhong, Lynn, & Hamel, 2000). Much of the best research to date is from abroad. In England, for example, Hinton (1999) performed weekly interviews with dying patients, monitoring changes in mood, anxiety, quality of life, and awareness and acceptance of death over the last months of life. In Australia researchers have conducted the best longitudinal research on psychological adjustment among dying patients. Advanced melanoma and breast cancer patients were measured every 3 months for 2 years or until death. Their findings include trajectories of psychological adjustment (Brown, King, Butow, Dunn, & Coates, 2000) and coping (Brown, Brown, et al., 2000), as well as evidence that psychosocial factors predict length of survival among patients with advanced cancer (Brown, Butow, Culjak, Coates, & Dunn, 2000; Butow, Coates, & Dunn, 1999, 2000). Despite these fascinating studies, psychological states require much additional attention. Spiritual concerns have received even less attention, despite the fact that qualitative research suggests that they are broadly endorsed as important for quality of dying (e.g., Barnard et al., 2000; Ehman, Ott, Short, Ciampa, & Hansen-Flaschen, 1999). Again, longitudinal studies are needed rather than cross-sectional ones. Many of us have “common sense” expectations about spiritual issues and dying, for example, that religious beliefs become more salient in the face of mortality. Inclusion of spiritual issues in studies of dying would generate important knowledge about these issues.

Mental Health Issues in Terminal/Palliative Care.—Another gap in end-of-life research is lack of knowledge about the prevalence of diagnosable mental illness among dying patients and the effects of treating versus not treating mental illness during the last weeks or months of life. In the few studies available, estimates of major depression among terminally ill patients ranged from 17% to 42% (Bukberg, Penman, & Holland, 1984; Heaven & Maguire, 1998; Le Fevre, Devereux, Smith, Lawrie, & Cornbleet, 1999). All the studies were restricted to patients with advanced cancer, and the sample sizes were small, ranging from 62 to 87. Similarly, the prevalence of psychotropic drug use by dying patients is unknown. If one of the goals of end-of-life care is to meet the needs of the whole person, psychological as well as physical, there must be a knowledge base about the prevalence and correlates of mental illness and about treatment options and their efficacy.

The Role of Comorbidities.—The importance of measuring comorbidities in studies of the health and functioning of older adults is well-documented. Surprisingly, virtually no attention has been paid to comorbidities in end-of-life research. In the single study found that included comorbid diagnoses, the number of conditions significantly contributed to symptom burden (physical and emotional) among terminally ill patients (Desbiens, Mueller-Rizner, Connors, Wenger, & Lynn, 1999). More typically, study participants are identified as cancer patients or stroke victims or in terms of other single diagnoses. It is understandable that attention would focus in large part on the presumably fatal diagnosis. Nonetheless, the trajectory of dying surely differs among patients with the same fatal diagnosis, but who differ in terms of the number and severity of other morbidities, pretreatment functional status, and general frailty. Presumably, these factors affect treatment decisions both before and during the process of dying. They also may have important implications for more behavioral or social facets of dying, such as choices for aggressive versus palliative treatment, advance directives, and even place of death.

Further Exploration of Diversity.—Although some useful studies exist, further research is needed to examine the ways that diversity is related to end-of-life issues. Research on age differences in end-of-life care and preferences has addressed both age differences within the older population (e.g., above and below a cutoff-point, typically 75 or 80 years of age) and, to a lesser extent, between older and younger patients. There is clear evidence that older patients, especially the very old, are less likely to receive specialty care, life-sustaining technology, or aggressive care of any kind and that they account for fewer admissions to the ICU (Auerbach et al., 2000; Hamel, Teno, et al., 1999; Hamel et al., 2000; Hanson & Danis, 1991). As a consequence, end-of-life health care expenditures are lower for elderly persons than others and for the oldest old than the young-old (Hamel et al., 1996; Temkin-Greener, Meiners, Petty, & Sztradowski, 1992). Note, however, that total costs are not the same as cost-effectiveness. There is general consensus that the cost-effectiveness of specific procedures depends on patient prognosis, severity, and comorbidities—not on age per se (e.g., Hamel et al., 2001). In addition to receiving less aggressive care at the end of life, older adults also are less likely to desire it (e.g., Phillips et al., 1996).

Examination of racial/ethnic differences in end-of-life preferences and care has been largely limited to comparisons of White and African American people. With regard to preferences, there is consistent evidence that African American persons are more likely than White persons to prefer aggressive treatment during the terminal phase of illness (this is observed behaviorally as well in terms of lower rates of do not resuscitate orders and advance directives; Blackhall et al., 1999; Garrett, Harris, Norburn, Patrick, & Danis, 1993; Hopp & Duffy, 2000; O’Brien et al., 1995). Evidence about race differences in the amounts and kinds of end-of-life care received are inconsistent.
Some investigators report that African American people receive more aggressive treatment than White people (e.g., Koch, Rodeffer, & Wears, 1994), whereas others find no race differences in use of life-sustaining technology (e.g., Phillips et al., 2000). These inconsistencies may reflect different policies at different health care settings, or they may result from methodological differences across studies (e.g., Koch and colleagues [1994] did not control on patient clinical characteristics and preferences; Phillips & coworkers [2000] did). There is strong evidence, however, from national Medicare beneficiary files that African American people have much higher end-of-life costs than other racial/ethnic groups (Hogan, Lunney, Gabel, & Lynn, 2001).

Although a sizable number of studies have examined age and race differences in end-of-life preferences, patterns of care, and expenditures, gaps in our understanding of these differences remain. First, more multivariate analyses are needed where potential confounding factors are statistically controlled. This is necessary before we can conclude with confidence that these differences are “real.” Second, if there are real age and ethnic differences, additional research is needed to identify the processes/mechanisms that generate those differences, especially if they reflect issues of ethics or social justice. These mechanisms are likely to be complex, involving culture, socioeconomic status, and cohort experiences. In addition, differences may emerge as a result of patient, family, provider, and/or even health system factors. Third, research is needed to determine whether these differences affect end-of-life outcomes, including length of survival, quality of life, and management of pain or symptoms. Borum, Lynn, and Zhong (2000) report Black-White differences in levels of intervention, with African Americans receiving less, but observed no race differences in length of survival or pain control, concluding that the differences were of “modest clinical significance.” This issue requires additional effort before that conclusion can be generalized beyond that study. Another high priority is to expand the study of racial/ethnic differences beyond White and African American people.

Provider Effects.—Although it is probably not an issue that physicians, nurses, and other health care professionals would welcome, research that assesses how much variance in patient outcomes can be explained by the individual provider (e.g., the attending physician) may be worth serious consideration. Research in other areas reinforces this recommendation. For example, educational research demonstrates that not all teachers are equally effective in motivating students and transmitting knowledge and skills (e.g., Erdle, Murray, & Rushton, 1985; Parikh, Patel, & Patel, 1984). It is possible that provider differences are important in end-of-life research as well. There may be important differences, for example, in multiple kinds of physician behavior, such as communication skills, time spent talking to the patient or the patient’s surrogate, and willingness to discuss end-of-life decisions. By the same token, consumer satisfaction, in part, may reflect reactions to the styles and behaviors of health care providers. If one element of high-quality, end-of-life care is compassionate and caring interactions with health care providers, it may be useful to determine the extent to which provider differences affect patient or family outcomes.

Examining Health System Variables.—Another relatively neglected issue is the potential impact of health system variables on end-of-life outcomes. For obvious reasons, a common assumption in end-of-life research has been that the “action” in end-of-life care is observed in the attitudes, preferences, and behaviors of terminally ill patients, their families, and their health care providers, especially their physicians. However, all of these persons are constrained to some degree by the health systems in which they are embedded. Two questions are how much are they embedded and what difference does it make? There are nascent indications that health system variables may play an important role in end-of-life care—not to the exclusion of patient, family, and provider factors, but in addition to them. Two recent studies provide initial evidence to buttress this recommendation. Using Medicare records and other national data, Christakis and Iwashyna (2000) examined a national cohort of more than 150,000 hospice patients until death. Although some personal characteristics of hospice patients were associated with earlier admission, local health system variables were significant predictors as well. Specifically, earlier hospice admission was observed for patients who resided in areas with more hospital beds, greater hospice capacity, and high ratios of general practitioners to specialists. In a different vein, Pritchard and colleagues (1998) found that rates of hospital deaths were higher in regions that had higher per capita hospital beds. This finding does not suggest that place of death is a simple combination of patient preferences and medical needs. It is methodologically and logistically difficult to build health system variables into many end-of-life studies; for example, when one samples only one health care facility, as is typical, health system factors are constants rather than variables. When possible, however, including health system variables in large-scale studies is recommended.

Moving Beyond Health Care to Supportive Services.—A final gap in end-of-life research is the nearly universal absence of information about services other than medical care that dying patients receive and their effects on patient/family outcomes and costs of terminal care. For persons who die at home after an illness of more than minimal duration, supportive services are often critically important elements of terminal care (e.g., Collins & Ogle, 1994; Duke, 1997). Home health workers, meals-on-wheels, visiting nurses, and other services are part of the professional network that serves terminally ill persons in their homes. Individuals who die in the in-patient hospice or a hospital often make extensive use of supportive services before...
their final admission to a health care setting (e.g., Karlsen & Addington-Hall, 1998). Nursing homes, of course, deliver supportive services, as well as medical care. To date, we know very little about how supportive services relate to different trajectories of dying or to patient or family outcomes. It is time for end-of-life research to broaden its scope and incorporate the full range of services received by terminally ill persons. Inclusion of supportive services also has important implications for studies of end-of-life expenditures. For example, neglect of community-based services, often not included in Medicare databases, artificially lowers the total costs associated with dying at home.

Conclusions

This review should not be viewed as an indictment of end-of-life research at this time. Much progress has been made on this important topic over a short time, and most of the problems that have been identified with end-of-life research can be observed in all "young" research fields. There is clearly much additional work to be done before we can be satisfied with the knowledge base of end-of-life research.

Both conceptual and methodological advances are high priorities for future research. Conceptually, the most important task is for the field to establish one or more definitions of dying that are used across studies. Other conceptual needs include (a) distinguishing among quality of life, quality of death, and quality of end-of-life care, (b) identifying appropriate outcomes in addition to consumer satisfaction, and (c) enlarging explanatory models to link preterminal states and characteristics with parameters of the dying process. Methodologically, the highest priority is longitudinal studies that cover longer segments of time than those in research to date.

References


Desbiens, N. A., Mueller-Rinzner, N., Connors, A. E., Jr., Wenger, N. S., & Lynn, J. (1999). The symptom burden of seriously ill hospitalized pa-


Ehman, J. W., Ott, B. B., Short, T. H., Ciampa, R. C., & Hansen-Flaschen, J. (1999). Do patients want physicians to inquire about their spiritual or religious beliefs if they become greatly ill? Archives of Internal Medicine, 159, 1803–1806.


Steinhauser, K. E., Clipp, E. C., McNeilly, M., Christakis, N. A., McIntyre, L. M., & Tulsky, J. A. (2000). In search of a good death: Observations of patients, families, and providers. *Annals of Internal Medicine, 132*, 825–832.


Received September 6, 2001
Accepted June 19, 2002