Abstract

Context. There has been surprisingly little attention to conceptual and methodological issues that influence the measurement of discretionary utilization at the end of life (DIAL), an indicator of quality care.

Objective. The objectives of this study were to examine how DIALs have been operationally defined and identify areas where evidence is biased or inadequate to inform practice.

Methods. We conducted a scoping review of the English language literature published from 1/1/04 to 6/30/17. Articles were eligible if they reported data on ≥2 DIALs within 100 days of the deaths of adults aged ≥18 years. We explored the influence of research design on how researchers measure DIALs and whether they examine demographic correlates of DIALs. Other potential biases and influences were explored.

Results. We extracted data from 254 articles published in 79 journals covering research conducted in 29 countries, mostly focused on cancer care (69.1%). More than 100 DIALs have been examined. Relatively crude, simple variables (e.g., intensive care unit admissions [56.9% of studies], chemotherapy [50.8%], palliative care [40.0%]) have been studied more frequently than complex variables (e.g., burdensome transitions; 7.3%). We found considerable variation in the assessment of DIALs, illustrating the role of research design, professional norms and disciplinary habit. Variables are typically chosen with little input from the public (including patients or caregivers) and clinicians. Fewer than half of the studies examined age (44.6%), gender (37.3%), race (26.5%), or socioeconomic (18.5%) correlates of DIALs.

Conclusion. Unwarranted variation in DIAL assessments raises difficult questions concerning how DIALs are defined, by whom, and why. We recommend several strategies for improving DIAL assessments. Improved metrics could be used by the public, patients, caregivers, clinicians, researchers, hospitals, health systems, payers, governments, and others to evaluate and improve end-of-life care. J Pain Symptom Manage 2020;59:894–915. © 2019 The Authors. Published by Elsevier Inc. on behalf of American Academy of Hospice and Palliative Medicine. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Key Words
Aggressive care, aging, biomedical ethics, dying, end-of-life care, health care quality, hospice, ICU, life-prolonging, palliative care
Introduction

End-of-life (EoL) care is an increasingly contested arena, as the industrialization of medicine, rules of rescue, medical assistance in dying, and health care costs are debated in professional journals, the popular press, and legislative bodies. Nowhere is this conflict more fraught than in the delivery of interventions shortly before patients die. Even the language used to describe EoL care is contested. Most individuals say they prefer to die peacefully at home, but many die in hospitals shortly after receiving treatments that have been described using value-laden terms, such as potentially avoidable, burdensome, inappropriate, intensive, unduly intensive, futile, wasteful, and aggressive.

In this study, we use the neutral term discretionary to refer to interventions that typically confer benefits to patients but have limited or no evidence for their medical efficacy in particular clinical contexts. Offered at the discretion of clinicians, patients, and family members after weighing the benefits and burdens of their use, discretionary interventions, treatments, procedures, or services can cause more harm than good if underused or overused. For example, intubation and mechanical ventilation are well-established medically beneficial interventions for respiratory failure in the context of severe pneumonia that allow clinicians time to treat and resolve the underlying pneumonia. However, those same interventions are rarely medically beneficial among patients with advanced cancer and comorbid pneumonia, even when they confer nonmedical (i.e., social, emotional) benefits that are important to some patients and family members. Thus, the same intervention can be either standard or discretionary depending on the clinical circumstances.

Whereas the debates about EoL care reflect remarkable sensitivity to clinical subtleties and language use, there has been surprisingly little attention to the nuanced measurement issues that influence the interpretation of data on discretionary utilization at the end of life (DIAL). Beginning with Earle’s seminal papers on DIAL measurement in cancer care, hundreds of articles reporting data on DIALs have appeared, and efforts are underway to develop quality metrics. Several impactful papers outlining methodological issues in EoL research have been published, but we are aware of no prior papers that focused exclusively on DIAL measurement.

In an effort to improve the transparency of DIAL measurement and mitigate demographic disparities in EoL care, this scoping review will answer the following broad questions: how have DIALs been operationally defined in peer-reviewed research? Does research design influence how researchers measure DIALs, or whether they examine demographic correlates of DIALs? Is there other evidence of bias in the literature? Whereas meta-analyses and systematic reviews answer focused questions of a narrowly defined, mature literature, scoping reviews aim to answer broad questions of an emerging literature that is dispersed across multiple disciplines and sub-specialties. Unlike meta-analyses or systematic reviews, scoping reviews have a broad purview and can explore the scientific process itself, by, for example, critically analyzing the manner in which research questions have been generated, key variables have been operationalized, and knowledge has been created. Meta-analyses and systematic reviews can have immediate impact for practice but scoping reviews motivate and advance scholarship by identifying areas where concepts are underdeveloped and evidence is biased, absent, or inadequate to inform practice.

This scoping review will provide a map to those interested in foraying into the poorly charted terrain of DIAL measurement. Its timeliness is underscored by the clinical and policy challenges posed by two demographic trends, the aging of the large baby-boom cohort and the growth in populations worldwide of patients with multiple comorbidities at risk for receiving DIALs. DIALs have been shown to be associated with worse patient quality of life before death, worse bereavement outcomes in personal caregivers, elevated health care costs, missed opportunities to divert precious resources elsewhere, clinician burnout, and staff turnover. Without identifying the conceptual and methodological traps that undermine reliable, valid, credible measurement of DIALs, it will be difficult to improve patients’ experiences at the EoL or validly evaluate systems of care designed to improve utilization and other important outcomes. Similarly, it will be difficult to ascertain whether EoL care is improving over time or getting worse, better in some regions than others, or worse in some demographic groups than others.

Methods

Information Sources

The literature search was implemented primarily in PubMed as well as PsycINFO, Web of Science, and Google Scholar. After experimenting with combinations of keywords (aggressive, end-of-life, futile, health care utilization, health services, intensity, intensive care, older adult, quality, quality indicators, quality of care, terminal, treatment), we settled on the following: (end-of-life OR end of life) AND quality AND (hospice OR palliative OR invasive OR life-prolonging OR intensive OR intensity OR aggressive). The reference sections of several review articles were examined (so-called backward searches).
Forward searches were conducted, the web sites of governmental organizations (e.g., AHRQ, NCI, Medicare) were examined, and nontraditional search methods were used (e.g., Google recommended article function). No journals were hand-searched.

**Inclusion Criteria**

We included English language articles published in 2004 or later (until June 30, 2017) that reported quantitative data on EoL interventions, treatments, and procedures. The last search date was September 30, 2017. We chose to begin in 2004 because a seminal paper was published in 2003.33 (That paper was itself motivated by a 2000 Institute of Medicine report.68) We excluded review articles, commentaries, and qualitative or mixed-methods reports without quantitative data on specific indicators of discretionary utilization. Economic analyses were included if data were provided on the costs of named components of EoL care (e.g., hospice). Three other eligibility criteria were required:

**Reported on Two or More DIALs.** Our starting assumption is that discretionary utilization at the end of life is a latent construct,69,70 meaning that DIALs cannot be precisely measured in the same sense that one can measure physical attributes such as height, weight, or body temperature. Like other latent constructs, such as intelligence, depression, or consumer prices, DIALs can only be estimated or approximated using multiple (two or more) indicators that tend to be correlated.31,71-75 Therefore, we excluded articles that reported on only one indicator of DIALs, such as chemotherapy,76-78 intensive care unit (ICU) stays,79 place of death,80,81 or hospice.81

**Reported Data on Adult Patients Presumed to Have Capacity to Make Medical Decisions.** We excluded studies of children (<18 years of age) and persons with dementia.82 Children rarely have legal capacity to make treatment decisions. Dementia is known to influence how EoL decisions are made. Patients typically take a less active role and receive less aggressive treatment than persons without dementia.82 We considered confining the review to studies of participants who were cognitively intact, but few studies reported relevant data.

**Reported Data on DIALs Within 100 Days of Death.** Determining when the EoL begins—either biologically (e.g., a “biomarker”) or temporally (e.g., days before death)—remains elusive. Given the unavailability of a reliable biomarker demarcating the EoL, that is equally valid in persons with different diagnoses, we used a temporal marker. Recognizing that transitions from seriously ill to actively dying are rarely knowable,84 and that any temporal cut point is arbitrary, the review was confined to studies that reported data on the final 100 days. It is easier to predict death within 100 days than within 6 months or longer.85,86 In contrast to longer timeframes, the 100-day timeframe permitted us to examine DIALs before deaths that are more predictable.85-89 Studies reporting data more than 100 days before death were included if they also reported data within the final 100 days.

**Data Abstraction**

The first author worked with a librarian and another author (M. C.) to screen the search results (titles, abstracts) against the eligibility criteria. Potentially relevant full-text articles were prescreened. Discrepancies across screeners were resolved through discussion.

The data abstraction form (Appendix I) was iteratively modified beginning in July 2016 and finalized in April 2017. The first author trained all coders (undergraduate students, graduate students, research assistants) to use the form. Coders were allowed to independently abstract data after they demonstrated they could code with a high degree of accuracy (fewer than five errors). To minimize error, data were extracted by at least two coders, with small teams focusing on particular variables. The first author resolved all coding discrepancies in consultation with a dually trained nephrologist/palliative care physician (F. S.) or a nurse with expertise in palliative care and intensive care (S. N.). The first author personally checked the accuracy of all coded variables on all forms.

**DIALs (Presence/Timing).** The abstraction form contained 13 prespecified DIALs (closed-ended, dichotomous items) as well as open-ended items. Closed-ended items forced coders to determine whether articles reported data on the prespecified DIALs. Largely reflecting Earle’s research in medical oncology85,54,73 and Barnato’s studies of critical care,31 the prespecified DIALs were as follows: inpatient hospitalization, emergency department (ED) visits, admission to an ICU, chemotherapy, cardiopulmonary resuscitation, mechanical ventilation, dialysis/hemodialysis, intubation, tracheostomy, artificial nutrition (e.g., tube feeding, parenteral nutrition), prescription medication (e.g., antibiotics, vasopressors), site of death, hospice and palliative care. As the terms palliative care and hospice do not have universally shared meanings, the words “palliative care” and “hospice” had to appear in the article to be coded. Palliative chemotherapy was coded as chemotherapy.

The presence of open-ended items enabled coders to identify DIALs that were not prespecified. Data on the number of days before death covered by the DIAL assessment (one, two, three, seven, 14, 30, 60, >60 days before death) were coded (e.g., inpatient hospitalization seven days before death).
Sample Characteristics. For each included article, we coded the number of decedents, diagnostic population (individuals with cancer, with another specific diagnosis [e.g., renal disease], or other), age group (≥65, unknown, other), gender composition (% men), and country of data collection.

Study Design. Latent constructs (e.g., DIALs, consumer prices) can only be estimated using multiple indicators, not measured with near-pinpoint precision as in the measurement of height. Consequently, this leaves room for unwarranted, systematic variation in estimates, reflecting the vagaries of measurement. Methods for deriving DIALs estimates are subject to biases related to investigator attributes (motives, disciplinary norms, etc.), estimation methods (e.g., instrumentation), or study design. To examine the effects of study design on seemingly unwarranted variation, or bias, in DIALs assessments, studies were coded as prospective, retrospective, or terminal hospitalization. The latter focuses on a sample of hospitalized patients, all of whom died while hospitalized. Because these studies report data on a sample of decedents, they could be classified as retrospective, but they could also be classified as prospective because they use data collected prospectively and documented in a local medical record. Documentation could, in theory, affect EoL care, analogous to the Hawthorne effect.90 We thus created a separate category to identify studies of terminal hospitalization.

Sociodemographic Correlates. We determined whether articles reported demographic (age, gender, race) or socioeconomic (annual household income, occupation, education) correlates of DIALs. Despite justifiable, ethical91 concerns about demographic disparities in EoL care in the U.S.,64,65,92 and elsewhere,93–96 this is the first attempt to explore the effect of study design on the examination of demographic differences in EoL care.

Publication Characteristics. We recorded journal names in an effort to gauge interest in DIALs outside of palliative care and oncology as well as to identify patterns of DIAL assessments as a function of medical specialty. Given the need for timely reporting of findings, we created a timeliness index by computing the temporal lag between the end of data collection (month/year) and date of initial (print or online) publication. To determine if authors justify their decisions to study particular DIALs (e.g., ED visits) over a particular timeframe (e.g., final 7 days of life), we coded whether they offered a citation to justify their methodology. Given our interest in identifying biasing influences on DIAL assessments, we attempted to reliably code author’s motives for conducting the research (e.g., demonstrate positive impact of palliative care, test hypotheses about EoL interventions, document harms of overtreatment).

Synthesis

Synthesis occurred in three overlapping stages. The first stage occurred as data were being abstracted, coded, and entered. When coders were uncertain whether an article reported a particular DIAL, they were instructed to discuss the article with a supervisor (P. D. or M. H.). Final decisions were made by the first author in collaboration with a nurse and/or a physician.

In the second stage, data were statistically analyzed and then presented and discussed in a series of seven 60- to 90-minute meetings (6/8/17 to 8/31/17) attended by coders and a multidisciplinary team, representing a range of disciplines and specialties (palliative care, nursing, oncology, nephrology, internal medicine, family medicine, psychology). In these meetings, quantitative summary data were presented along with inferential statistics comparing prospective, retrospective, and terminal hospitalization designs. Approaches to categorizing and tabulating the data were developed, emerging themes were identified, and preliminary drafts of this manuscript were reviewed. In the third stage, the manuscript was iteratively edited by all authors and was critiqued by two palliative care scholars who were not involved in the project’s earlier stages and did not co-author this paper. It was modified again after receiving input from five anonymous reviewers.

Results

General Overview

After screening more than 5300 abstracts, the investigators read more than 400 articles, nearly half of which were excluded because the article reported data on only one intervention (e.g., chemotherapy use; see Section Reported on Two or More DIALs); provided data on nondecedents and decedents, without distinguishing the two (because our focus is on decedents); provided no information about the timing of the assessment in relation to the date of death (because we focused on DIALs in the final 100 days of life); or solely reported economic (cost) analyses. This left us with 254 articles (Fig. 1) over the course of the study period (1/1/04 to 6/30/17), with most (n = 164, 64.6%) published since 2013 (Fig. 2, see Appendix II for included articles). Six articles31,74,97–100 reported data on two types of cohorts (e.g., prospective, terminal hospitalized). The total number of unique cohorts examined was thus 260.
Studies have appeared in 79 different journals, ranging from general medicine (New England Journal of Medicine, Lancet), surgery (Journal of Surgical Research), and critical care (American Journal of Respiratory and Critical Care Medicine) journals to health disparities (Journal of Healthcare for the Poor and Underserved), health services (Medical Care), health policy (Health Affairs), and gerontology (Gerontologist) journals. Most articles appeared in cancer (96 articles) or palliative care (73 articles) journals, particularly the Journal of Palliative Medicine (27), Journal of Pain and Symptom Management (17), Cancer (17), Supportive Care in Cancer (17), and Journal of Clinical Oncology (16). Outside of oncology and palliative care, the most common outlets were the Journal of the American Geriatrics Society (8), JAMA Internal Medicine (8), and JAMA (7).

Studies have been conducted in 29 different countries, mainly in the U.S. (137 studies), Canada (26), The Netherlands (16), Italy (13), and Taiwan (13). With the exceptions of articles from Turkey, Lebanon, Brazil, and the People’s Republic of China, studies have originated from countries with advanced economies, as defined by the International Monetary Fund (www.imf.org).

Fig. 1. PRISMA flow diagram for scoping articles. aThis number and subsequent numbers are approximations, as traditional (backward search, forward search) and nontraditional (e.g., the Google “recommended articles” function) methods were used to identify additional records, and the number of articles generated using nontraditional methods were not tallied.

bSix articles reported data using more than one sample and research design; the total number of samples was 260. See text for details.
Table 1 shows that most articles (180; 69.3%) focused on cancer; the remainder examined persons with another disease, nursing home residents, patients in primary care practices, or community samples. The median (IQR) number of deaths reported was 639.5 (265, 8148). Men constituted 55.1% (±19.0%) of the samples.

Comparing Prospective, Retrospective, and Terminal Hospitalization Designs
The literature is dominated by retrospective studies (Table 1). These studies had larger sample sizes, were more likely to focus on older adults, and were more likely to use administrative data sets. Prospective articles reported more timely data than other articles. The [median (IQR)] number of months from last death observed to date of publication was 27.5 (17, 55) in prospective studies and 37 (25, 58) in terminal hospitalization studies (z = 3.04, P < 0.0024, nonparametric Wilcoxon rank-sum test).

The most commonly studied variable was site of death, which was examined in 73% of articles that used a retrospective or prospective design. Table 2 shows that, after site of death, the most frequently examined DIALs were ICU stays, inpatient admissions, and chemotherapy.

We examined the effects of study design on seemingly unwarranted variation, or bias, in the ways DIALs have been operationally defined. Retrospective studies were more likely to report data on ED visits but less likely to report data on cardiopulmonary resuscitation and ventilation/intubation. Terminal hospitalization studies were more likely to report data on dialysis and medication use. Prospective studies were more likely to report data on the final month of life, prospective studies were more likely to examine the final week, and terminal hospitalization studies were more likely to focus on the final day (Appendix III).

We also examined the effects of study design on seemingly unwarranted variation, or bias, in the examination of demographic correlates of DIALs. Table 3 shows that retrospective studies were more likely to

<table>
<thead>
<tr>
<th>Variable</th>
<th>167 Retrospective Studies</th>
<th>46 Prospective Studies</th>
<th>47 TH Studies</th>
<th>Total (k = 260)</th>
</tr>
</thead>
</table>
| Number of deaths              | 2074 (319, 32,800)
- Focused on cancer patients only | 122 (73.1)           | 38 (82.6)                | 20 (42.6)        | 180 (69.3)               |
| Focused on older patients only | 51 (30.5)           | 5 (10.9)                 | 5 (10.6)        | 61 (23.5)               |
| Used administrative data set  | 104 (62.3)          | 9 (19.0)                 | 11 (23.4)       | 124 (47.7)             |

TH = terminal hospitalization; IQR = interquartile range.

| TH = terminal hospitalization; IQR = interquartile range. Bolded values in a column are significantly different from the values in the other two columns. *z = -5.33; P < 0.0001, nonparametric Wilcoxon rank-sum, $^1$Pearson chi² = 17.94, df = 1, P < 0.001, $^2$Pearson chi² = 12.92, df = 1, P < 0.001, $^3$Pearson chi² = 39.80, df = 1, P < 0.001. $^4$Most (38/46, 82.6%) articles in this category originated from the U.S.; 18 articles used one data set, Coping with Cancer (H.G. Prigerson, PI).

Fig. 2. Growth of the literature on DIALs, 2004–2016. There was a notable surge in interest in this topic in 2013, and that was not confined to palliative care (PC) and oncology (Onc) journals. The review included articles published before July 1, 2017; for consistency, 2017 data are not shown.
Table 2
Number (%) of Studies Reporting Data on 12 DIAL Indicators<sup>a</sup>

<table>
<thead>
<tr>
<th>Indicator</th>
<th>167 Retrospective Studies</th>
<th>46 Prospective Studies</th>
<th>47 TH Studies</th>
<th>Total (k = 260)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICU stay</td>
<td>88 (52.7)</td>
<td>31 (67.4)</td>
<td>29 (61.7)</td>
<td>148 (56.9)</td>
</tr>
<tr>
<td>Inpatient admission</td>
<td>112 (67.1)</td>
<td>20 (43.5)</td>
<td>10 (21.3)</td>
<td>142 (54.6)</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>91 (54.5)</td>
<td>25 (54.3)</td>
<td>16 (34.0)</td>
<td>132 (50.7)</td>
</tr>
<tr>
<td>Hospice</td>
<td>83 (49.7)</td>
<td>32 (69.6)</td>
<td>7 (14.9)</td>
<td>122 (46.9)</td>
</tr>
<tr>
<td>Palliative care&lt;sup&gt;g&lt;/sup&gt;</td>
<td>66 (39.5)</td>
<td>10 (21.7)</td>
<td>28 (59.6)</td>
<td>104 (40.0)</td>
</tr>
<tr>
<td>Emergency department use</td>
<td>86 (51.5)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>11 (23.9)</td>
<td>5 (10.6)</td>
<td>102 (39.2)</td>
</tr>
<tr>
<td>Ventilation/intubation</td>
<td>33 (19.8)&lt;sup&gt;f&lt;/sup&gt;</td>
<td>27 (58.7)</td>
<td>25 (53.2)</td>
<td>85 (32.7)</td>
</tr>
<tr>
<td>Cardiopulmonary resuscitation</td>
<td>24 (14.4)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>24 (52.2)</td>
<td>24 (51.1)</td>
<td>72 (27.7)</td>
</tr>
<tr>
<td>Feeding tube</td>
<td>23 (13.8)</td>
<td>14 (30.4)</td>
<td>20 (42.6)</td>
<td>57 (21.9)</td>
</tr>
<tr>
<td>Medications&lt;sup&gt;a&lt;/sup&gt;</td>
<td>20 (12.0)</td>
<td>5 (10.9)</td>
<td>20 (42.6)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>45 (17.3)</td>
</tr>
<tr>
<td>Dialysis</td>
<td>15 (9.0)</td>
<td>5 (10.9)</td>
<td>17 (36.2)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>37 (14.2)</td>
</tr>
<tr>
<td>Tracheostomy</td>
<td>4 (2.4)</td>
<td>3 (6.5)</td>
<td>8 (17.0)</td>
<td>15 (5.8)</td>
</tr>
</tbody>
</table>

DIAL = discretionary utilization at the end of life; TH = terminal hospitalization; ICU = intensive care unit; ED = emergency department.

Bolded values in a column are significantly different from the values in the other two columns; <sup>a</sup>Pearson chi<sup>2</sup> = 21.64, df = 1, P < 0.001; <sup>b</sup>Pearson chi<sup>2</sup> = 35.48, df = 1, P < 0.001; <sup>c</sup>Pearson chi<sup>2</sup> = 41.38, df = 1, P < 0.001; <sup>d</sup>Pearson chi<sup>2</sup> = 24.41, df = 1, P < 0.001; <sup>e</sup>Pearson chi<sup>2</sup> = 21.65, df = 1, P < 0.001.

1. We did not report data on site of death in the table because, by definition, all decedents in TH studies died in hospital. See text for data on site of death.

2. Although palliative chemotherapy, palliative radiation, and palliative surgery are part of palliative care in a broad sense, they are also arguably invasive and thus were not considered palliative care for the purposes of this review. Our assumption is that the undertuse of palliative interventions is inappropriate, but it is plausible that some palliative interventions are overused—just as hospice can be used inappropriately.

The most commonly studied medications were pain medications, antibiotics, vasopressors, and psychotropics.

Emergent Themes: Conceptualizing and Measuring DIALs

As the research group discussed the findings in a series of meetings, and the manuscript was revised, five themes emerged.

1. Variables differ considerably in complexity, ranging from simple dichotomies with uncertain clinical relevance (e.g., received chemotherapy in the final 30 days) to more complex, clinically relevant variables (e.g., received goal-concordant care). As shown in Table 4, it is easier and cheaper for researchers to acquire simple variables, but their simplicity compromises clinical relevance. In the simple approach, what counts, and is counted, is whether patients were hospitalized, on chemotherapy, visited an ED, died in the hospital, and so on. Contextual nuance is rarely considered. Simple variables examine patient exposure to treatments, procedures, or interventions, not clinical appropriateness (or avoidability) for a particular patient at a particular moment.

Reflecting dissatisfaction with simple variables, researchers have created complex variables (Table 5) in retrospective, prospective, and terminal hospitalization studies. These variables have been created using interview data, medical records, and administrative data to capture the avoidability of hospitalizations<sup>1</sup>; ED visits for chemotherapy toxicity<sup>74</sup>; the number of transfers and transitions across care settings<sup>12,13,100,107–110</sup>; hospital deaths of nursing home residents within three<sup>101</sup> or eight<sup>111</sup> days of hospital transfer; the deaths of terminally ill patients who had been transferred to the ICU and offered hemodialysis, vasopressors, or intubation<sup>99</sup>; ICU deaths of patients with cancer after resuscitation and/or mechanical ventilation; and receipt of goal-concordant care.<sup>102,112</sup>

2. Motives for reporting data are not transparent, raising questions about bias. Our attempt to discern researchers’ motives for reporting data

Table 3
Number (%) of Articles Reporting Demographic Correlates of DIALs<sup>a</sup>

<table>
<thead>
<tr>
<th>Examined Effects of:</th>
<th>167 Retrospective Studies</th>
<th>46 Prospective Studies</th>
<th>47 TH Studies</th>
<th>Total (k = 260)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>85 (50.9)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>12 (26.1)</td>
<td>19 (40.4)</td>
<td>116 (44.6)</td>
</tr>
<tr>
<td>Gender</td>
<td>73 (43.7)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>11 (23.9)</td>
<td>13 (27.7)</td>
<td>97 (37.5)</td>
</tr>
<tr>
<td>Race</td>
<td>40 (24.0)</td>
<td>18 (39.1)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>11 (23.4)</td>
<td>69 (26.5)</td>
</tr>
<tr>
<td>Socioeconomic status</td>
<td>40 (24.0)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>7 (15.2)</td>
<td>1 (2.1)</td>
<td>48 (18.5)</td>
</tr>
<tr>
<td>≥1 Demographic variable</td>
<td>96 (57.7)</td>
<td>19 (41.3)</td>
<td>26 (55.5)</td>
<td>141 (54.2)</td>
</tr>
</tbody>
</table>

DIAL = discretionary utilization at the end of life.

Bolded values in a column are significantly different from the values in the other two columns. <sup>a</sup>Pearson chi<sup>2</sup> = 7.46, df = 1, P = 0.006; <sup>b</sup>Pearson chi<sup>2</sup> = 8.19, df = 1, P = 0.004; <sup>c</sup>Pearson chi<sup>2</sup> = 4.55, df = 1, P = 0.033 [11 of these articles used data from Prigerson’s Coping with Cancer study]; <sup>d</sup>Pearson chi<sup>2</sup> = 9.35, df = 1, P = 0.002.

As shown in Table 4, it is easier and cheaper for researchers to acquire simple variables, but their simplicity compromises clinical relevance. In the simple approach, what counts, and is counted, is whether patients were hospitalized, on chemotherapy, visited an ED, died in the hospital, and so on. Contextual nuance is rarely considered. Simple variables examine patient exposure to treatments, procedures, or interventions, not clinical appropriateness (or avoidability) for a particular patient at a particular moment.
Variables differ considerably in complexity, ranging from simple dichotomies with uncertain clinical relevance (e.g., received chemotherapy) to more complex, clinically relevant variables (e.g., received goal-concordant care). Simple and complex variables also differ in availability and the extent to which they have served as outcomes in clinical trials. Some constituencies have prioritized simple variables; others are more invested in complex variables.

Some research appears to have been motivated by the need of specialists to identify specialty-specific quality indicators, raising questions about bias. Many variables appear to have been chosen on the basis of data availability and input from medical specialists who have a shared understanding of that specialty’s norms. Professional societies, through their publications and committees, reinforce this shared understanding. Studies conducted by medical oncologists that appear in oncology journals typically report data on chemotherapy use before death; studies conducted by nephrologists that are published in nephrology journals report on dialysis initiation; studies conducted by surgeons that appear in surgery journals report data on surgical procedures; studies conducted by intensivists that appear in critical care journals report data on ICU length of stay. This publication pattern is understandable, given the varied clinical interests of specialists and subspecialists and the need to identify quality indicators that reflect care processes over which specialists can exert control. Unfortunately, the focus of different constituencies (e.g., the membership of specialty professional organizations, payers, individual patients) is rarely well aligned. As shown in Table 4, some constituencies favor simple variables, and others favor complex variables. This misalignment represents a methodological challenge that could lead to metrics with uncertain relevance to the evaluation of overall quality care and the patient experience.

In an effort to illustrate a potential misalignment and critically analyze the manner in which key variables have been operationalized in the DIAL literature, we reviewed the logic underlying Earle et al.’s decision in 2008 to omit ED visits from their metric. We chose to discuss this metric because of its perceived importance. It was endorsed by the National Quality Forum, Center for Medicaid and Medicare Services, and the Agency for Health Research and Quality.

Earle et al. originally included ED visits in their metric. Five years later, Earle et al. abandoned that indicator, and others, arguing that the omitted indicators are “strongly influenced by comorbidity and, therefore, appear less useful as measures of aggressive cancer care” (pg. 3862). Although risk for ED use would seem to be relevant to oncologists’ decision making, and statistical adjustments for comorbidity are routinely performed, the authors reasoned that the newer metric would provide a more “methodologically sound” (pg. 3860) assessment of the work of medical oncologists. Earle et al. were appropriately circumspect about their methodology. They sought to identify variables reflective of an oncologist’s job performance, as viewed through a particular evaluative frame. They solicited patient and caregiver input for their 2003 paper, recognized that administrative data provide little information about treatment intent and could underestimate treatment aggressiveness. They identified variables that were unavailable in administrative data (advance directives) but could eventually be included in metrics when available.

Stakeholders were involved in developing the measure described in the 2003 paper, but we uncovered no evidence that caregivers or patients were consulted before the decision to drop ED visits. Had Earle et al. adopted a broader view of the oncologist’s role or solicited patient and caregiver input when contemplating their revised metric, they may have retained ED use. That said, the decision to remove ED visits from the
<table>
<thead>
<tr>
<th>First Author, Year, Design</th>
<th>Country</th>
<th>Variable</th>
<th>Data Source</th>
<th>Population</th>
<th>Methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hui, 2009&lt;sup&gt;17&lt;/sup&gt;</td>
<td>Canada</td>
<td>Appropriateness of interventions</td>
<td>Medical records</td>
<td>Cancer inpatients</td>
<td>Researchers evaluated the appropriateness of blood work (e.g., electrolytes, liver function tests), microbiology, imaging, hormonal therapy, radiation therapy, enrollment in Phase I trial, chemoradiation, and surgery</td>
</tr>
<tr>
<td>Muni, 2011&lt;sup&gt;105&lt;/sup&gt;</td>
<td>United States</td>
<td>Dying in the setting of full support</td>
<td>Medical records</td>
<td>ICU deaths in 15 hospitals</td>
<td>Not withdrawing life-sustaining treatments</td>
</tr>
<tr>
<td>Wright, 2014&lt;sup&gt;102&lt;/sup&gt;</td>
<td>United States</td>
<td>Health care transitions</td>
<td>Claims</td>
<td>Individuals ≥ 66 years old diagnosed with ovarian cancer</td>
<td>Changes over the final three and 30 days of life in the institutional health care provider identification number based on Medicare billing data and hospice files</td>
</tr>
<tr>
<td>Hockley, 2010; Finucane, 2013&lt;sup&gt;101,106&lt;/sup&gt;</td>
<td>United Kingdom</td>
<td>Inappropriate hospital days/death</td>
<td>Medical records</td>
<td>Nursing home residents</td>
<td>Researchers reviewed medical records to determine if death of frail, “deteriorating” nursing home residents occurred within three days of hospitalization for pneumonia or dehydration</td>
</tr>
<tr>
<td>Loggers, 2009; 2013, and Wright, 2010&lt;sup&gt;15,20,157&lt;/sup&gt;</td>
<td>United States</td>
<td>Intensive or aggressive EoL care</td>
<td>Medical records, caregiver report</td>
<td>Cancer patients enrolled in cohort study</td>
<td>EoL care was considered intensive or aggressive if patients died in an ICU after receiving resuscitation and/or mechanical ventilation in the last week of life</td>
</tr>
<tr>
<td>Moscovici da Cruz, 2015&lt;sup&gt;74&lt;/sup&gt;</td>
<td>Brazil</td>
<td>Medical futility</td>
<td>Medical records</td>
<td>Cancer patients admitted to the ICU</td>
<td>Researchers reviewed medical charts to determine if terminally ill patients had been transferred to the ICU and started on hemodialysis, vasopressors, or mechanical ventilation</td>
</tr>
<tr>
<td>Daly, 2016&lt;sup&gt;11&lt;/sup&gt;</td>
<td>United States</td>
<td>Potentially avoidable admissions</td>
<td>Medical records</td>
<td>Cancer patients admitted to the ICU</td>
<td>Using a standardized assessment tool, a multidisciplinary panel</td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
<td>Country</td>
<td>Study Design</td>
<td>Outcome</td>
<td>Data Source</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>---------</td>
<td>-------------</td>
<td>---------</td>
<td>-------------</td>
</tr>
</tbody>
</table>
| Gozalo, 2011 | 2011 | United States | Retrospective | Potentially burdensome transitions | Claims | Nursing home residents  
Any transfer in the last three days of life, a lack of continuity of nursing home facilities before and after a hospitalization in the last 90 days of life, more than two hospitalizations for any reason or more than one hospitalization for pneumonia, urinary tract infection, dehydration, or sepsis in the final 90 days |
| Teno, 2013 | 2013 | United States | Retrospective | Potentially burdensome transitions | Claims | Fee-for-service Medicare beneficiaries  
Changes over the final three days of life in institutional health care provider identification number based on the Medicare billing data |
| Biola, 2010 | 2010 | United States | Retrospective | Receipt of preference-concordant care | Caregiver report, nursing staff report | Nursing home residents  
Hospitalization or receipt of antibiotics, CPR, or tube feeding against the patient’s wishes (per proxy report) |
| Mack, 2010 | 2010 | United States | Prospective | Receipt of preference-concordant care | Medical records, caregiver report | Cancer patients enrolled in cohort study  
Receipt of life-extending care or symptom-directed care at the EoL that was inconsistent with preferences assessed 125 days (median) before death |
| Pasman, 2013 | 2013 | Netherlands | Retrospective | Receipt of preference-concordant care | Caregiver report | Older adults and those with advanced directive  
Using a mortality follow-back design, researchers asked caregivers about the deceased person’s treatment preferences and whether those preferences were honored |
| Temkin-Greener, 2015 | 2015 | United States | Retrospective | Site of death after transfer | Claims | Nursing home residents  
Hospital deaths within eight days of nursing home transfer |

(Continued)
Table 5
Continued

<table>
<thead>
<tr>
<th>1st Author, Year, Design</th>
<th>Country</th>
<th>Variable</th>
<th>Data Source</th>
<th>Population</th>
<th>Methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Setoguchi, 2008&lt;sup&gt;74&lt;/sup&gt; Retrospective</td>
<td>United States</td>
<td>Toxicity-related ED visit or hospital admission</td>
<td>Claims</td>
<td>Cancer patients</td>
<td>Chemotherapy toxicity was coded as present if the diagnosis at the time of the ED visit or hospitalization included one or more of the following: fever or infection; neutropenia or thrombocytopenia; dehydration or electrolyte disorders; nausea, vomiting, or diarrhea; anemia; constitutional symptoms; deep venous thrombosis or pulmonary embolus; malnutrition.</td>
</tr>
<tr>
<td>Abarshi, 2010&lt;sup&gt;108&lt;/sup&gt; Retrospective</td>
<td>Netherlands</td>
<td>Trajectories across care settings</td>
<td>Physician report</td>
<td>Expected, nonsudden deaths</td>
<td>Researchers asked general practitioners to identify the settings where patients had received care and the duration in each. Settings were classified as homes, care homes, nursing homes, palliative care units, and hospitals.</td>
</tr>
<tr>
<td>Gielen, 2010&lt;sup&gt;107&lt;/sup&gt; Retrospective</td>
<td>Belgium</td>
<td>Transitions across care settings</td>
<td>Claims</td>
<td>Persons &gt;40 years of age</td>
<td>Researchers classified care settings as hospital, inpatient palliative care unit, care home, home with home care, and home without home care.</td>
</tr>
</tbody>
</table>

CPR = cardiopulmonary resuscitation; ED = emergency department; GP = general practitioner; ICU = intensive care unit.

<sup>a</sup>This review covers articles published from 1/1/04 to 6/30/17. We are aware of at least one recent article that used a complex variable as an end point in a clinical trial.<sup>151</sup>

<sup>b</sup>The authors<sup>12</sup> reasoned that more than one hospitalization for these conditions is equivalent to more than two hospitalizations for other conditions because these conditions are potentially manageable with “appropriate advance care planning, without the need for hospitalization” (pg. 1213).

<sup>c</sup>In 2004, Earle et al.<sup>34</sup> conducted secondary analyses using a similar approach, but they did not report the findings. They counted inpatient admissions or ED visits as measures of aggressiveness if the “diagnosis-related group or first-position International Classification of Diseases (ninth revision) code description mentioned the patient’s cancer or its treatment” (pg. 316).
metric simplified a methodology that could be used to evaluate oncologists, given a narrower view of their role. As we discuss later, this raises difficult questions about the nature of quality metrics and the ethical obligations of those charged with developing them. A key question for further research is how researchers’ biases have shaped not only the measurement of EoL care but the provision of EoL care as well.

4. **Beyond the 13 prespecified DIALs, researchers have studied more than 90 other activities.** We classified these variables into nine categories (Appendix IV). The most common were other disease treatments (e.g., radiotherapy), invasive diagnostics (e.g., colonoscopy), and general support/monitoring (e.g., blood draw).

Three qualifiers provide a more refined appreciation of Appendix IV. First, at this early stage of methodological development, we erred on the side of including variables that are not definitively discretionary. Many symptom-focused activities are presumed to be vital, not discretionary, for patients who are dying. For instance, the activities listed under “Fluid Removal” include common procedures which confer benefit and arguably little harm in patients with very short prognoses. There is little comparative effectiveness research on many of the identified activities, however, and some treatments (e.g., antibiotics) confer risks to patients who had not been directly exposed to the medication. The same may be true of some of the seemingly benign activities listed in Appendix IV. Second, some interventions are offered both to relieve symptoms as well as to demonstrate caring, compassion, and tenderness, or other “off-label” reasons. This scoping review uncovered no systematic research on off-label reasons for prescribing discretionary interventions. Several articles alluded to this phenomenon, however. One research group summed

---

**Fig. 3.** Longitudinal variation in the percentage of publications reporting each DIAL. Publications were ordered chronologically, and the z-test for proportions was used to compare the frequency with which each DIAL was reported in the early studies \((n = 84)\) versus the most recent studies \((n = 91)\). Relative to the early studies, reporting on inpatient admissions \((47.6\% \text{ vs. } 65.9\%, \ P = 0.016)\), palliative care \((29.8\% \text{ vs. } 48.4\%, \ P = 0.015)\), and radiotherapy \((4.8\% \text{ vs. } 18.7\%, \ P = 0.005)\) increased in recent studies, whereas reporting on ICU stays \((66.7\% \text{ vs. } 47.3\%, \ P = 0.008)\) and ventilation/intubation decreased \((40.5\% \text{ vs. } 16.5\%, \ P < 0.001)\). ICU = intensive care unit.
up this dilemma in a discussion of tube feeding dying nursing home residents: “Although tube feeding can be construed as more aggressive than careful hand-feeding, it may also be perceived as an indicator of caring and not ‘abandoning’ residents .... a (demonstration of) caring and not ‘abandoning’ residents.”81 (pp. 825).

5. **Operational definitions of DIALs have changed over time.** Although our review covered a relatively brief timeframe, time trends in the assessment of DIALs were evident. In 2004, Earle et al.34 noted that “palliative care measures of admission to hospice or length of hospice stay were not included in the composite measure because they do not themselves indicate aggressive treatment” (pg. 316). However, Earle et al. soon began reporting data on hospice utilization,73,74,117 and others followed suit.118–122 Other trends in the operationalization of DIALs were observed, perhaps reflecting changes in the perceived importance of certain outcome measures, developments in health care delivery, and the increased interest in DIALs in medical specialties and academic disciplines beyond palliative care and medical oncology. Some of these trends are shown in Fig. 3.

**Discussion**

This review documents sufficient interest in the measurement of DIALs to warrant the investment of further time, effort, and resources to refine this area of inquiry and develop more robust (reliable, valid, credible) EoL metrics. Research on DIALs has been conducted in more than 25 countries and published in outlets cutting across medical specialties (e.g., cardiology, critical care, dermatology nephrology, oncology, palliative care, surgery) and disciplines beyond clinical medicine (e.g., epidemiology, health services research, health policy). We found considerable variation in the assessment of DIALs as a function of study design. This variation is seemingly unwarranted and illustrates how the vagaries of measurement and instrumentation can bias the assessment of important outcomes. We also found that retrospective studies were more likely to report effects of patient age, gender, and socioeconomic status on DIALs. This is likely due the larger samples in retrospective studies. Prospective studies were more likely to examine race differences in DIALs, perhaps because data on race are more likely to be missing from clinical or administrative data sets.123 Based on the data synthesized here, we offer five prioritized recommendations to improve research guiding the development of EoL metrics (e.g., DIALs) that could be used by the public (including patients and caregivers), clinicians, researchers, hospitals, health systems, payers, governments, and nongovernmental organizations.

**The Future of DIAL Assessments**

**Motivate Public Involvement in the Development and Ongoing Revision of EoL Quality Metrics.** No fewer than 100 DIALs have been examined over the final 100 days of life. Many appear to have been chosen on the basis of input from medical specialists who have a shared understanding of that specialty’s norms and an interest in self-policing by developing quality standards. Patients, family members/caregivers, and other stakeholders have rarely been consulted, though there are exceptions.31,33,124

Developing, refining, and revising usable EoL care metrics for clinical trials, scorecards, and benchmarking requires the ongoing involvement of motivated citizens who understand the relevance of such metrics to their lives, local communities, and, some would argue, public health.125 We are not advocating for a zero-tolerance approach to the involvement of experts and medical societies in metric development, but we are calling for a more inclusive, democratic approach that is publicly accountable. By ensuring that specialists do not exert disproportionate influence in the development of quality indicators,125,126 the public interest can be prioritized. Parties with self-interest in a simple end point (e.g., chemotherapy use over the past 30 days) risk ignoring or discounting outcomes of greater relevance to other parties. This can lead to ethically difficult decisions.127

Forging sustained128,129 partnerships130 with citizen stakeholders and encouraging public deliberation about care of the dying and EoL metrics125,126,131,132 should improve metric robustness. If motivating public involvement in discussions about health care and policy is difficult, motivating public involvement in discussions about EoL metrics and discretionary interventions might seem impossible. Nonetheless, there have been promising developments and some cautionary tales. The success of the Conversation Project illustrates public interest in improving EoL care,133 but research on public attitudes toward low value treatments34 shows that about one-third of the public cannot readily envision deriving any benefit from avoiding discretionary treatments.

To motivate productive public discussions, researchers and policymakers might consider deliberative polling.135,136 This technique combines traditional public opinion polling with expert-led small group public deliberations. For example, small groups of citizens (including patients and family members) could be engaged in deliberative discussions to consider the items listed in Table 2, Table 5, and Appendix IV. How would they rank-order (or
prioritize) these indicators? Do they think health care systems or clinicians should be held accountable for failing to meet certain standards? How often do they think metrics should be revised? Why might there be a need for revision? What outcome variables should be used in clinical trials and tracked by governments, patient advocacy groups, and health care systems?

No single, definitive approach to EoL metric development (or EoL care) will be universally endorsed by all constituents. Without public input, experts and specialists with vested interests in the use of particular metrics, including corporations, will continue to exert disproportionate influence on matters of public policy and resource allocation. Of course, public involvement is not a panacea, and conflicts between an ethic that prioritizes individual life extension and an ethic of population well-being have historically been irreconcilable. Nonetheless, by engaging the public in discussions of DIALs, the effects of special interests could be minimized. Metrics proposed to guide policy and clinical care will have more public credibility.

**Develop Quality Metrics That Reflect the Complexity of How Clinicians Conceive of EoL Interventions.** Research has not properly reflected the realities of clinical practice, where interventions are offered for multiple reasons beyond disease treatment or symptom management, including, unfortunately, gaming the system. We are aware of no efforts to accommodate these realities in metric development.

To what extent do clinicians think they are offering interventions for curative reasons, palliative reasons, psychological reasons, simply because technologically intensive treatment is the norm in their institution, or some combination thereof? A cross-national survey of this topic could be illuminating, as would qualitative research. In addition, it will be important to understand the presence of variability in expert raters’ judgments of the appropriateness of health care interventions. For example, are intensivists more likely to view a hospitalization as potentially avoidable than oncologists or hospitalists because they are more likely to be exposed to medical futility and thus more likely to experience moral distress? By reflecting the multidetermined nature of EoL interventions, future quality metrics will provide a more genuine depiction of the care provided in clinical settings.

**Develop Patient-Centered DIAL Assessments for Use in Clinical Trials.** Beginning with SUPPORT, many interventions studied in clinical trials have attempted to enhance the quality of EoL care. Nonetheless, utilization outcomes in many trials have been examined using simple variables, with debatable relevance to clinical care. A dying young woman might wish to be transferred to the ICU and placed on a ventilator to have an opportunity to hold her son on his first birthday. A dying middle-aged man might opt for another round of chemotherapy and to begin dialysis, just to be able to attend his daughter’s wedding. Viewed from one decision frame, a clinician who starts an invasive treatment while patients are dying might be seen as acting unethically. Viewed differently, the clinician’s behavior might be construed as empathic and moral. Without data on patient goals or the perspectives of family members or caregivers, the clinician’s actions are uninterpretable.

Only recently have researchers begun to attempt to assess the receipt of goal-concordant care, and many challenges remain. The first clinical trial to report data on the concordance of patient-preference and actual care received at the EoL was published in 2019 and was thus excluded from this review. A recently published observational study was similarly omitted. These early studies point to the need for further methodological and conceptual development. In particular, there is a need for mixed method research that links quantitative utilization data with data that provide contextual nuance about patients’ and family members’ shifting goals and values.

**Continue Developing “Basket-of-Goods” Approaches to DIAL Assessments.** Beyond evaluating the quality of care provided to individual patients enrolled in trials or seen in particular care settings, there is a need to develop indices that could be used to evaluate structural problems across entire health systems or in larger units of analysis (e.g., hospital referral regions; cities, counties, etc.). This is best accomplished using a “basket-of-goods” approach, analogous to those used to evaluate consumer prices. Although more than 100 variables have been used to assess DIALs, most articles report data on just two or three, such as site of death or chemotherapy. This is like estimating consumer prices or inflation without including housing, transportation, or medical costs. A more diverse basket-of-goods index approach to DIAL measurement will provide a truer reflection of EoL care while also making it more difficult for payees to manipulate metrics to generate revenue. Efforts to create a basket-of-goods DIAL index are currently underway. Building on research conducted for the Dartmouth Atlas, the Coalition to Transform Advanced Care recently unveiled the Advanced Care Transformation Index℠ (ACT Index℠), a 25-item index developed with some public input.

**Enhance the Manuscript Peer-Review Process to Improve the Timeliness and Transparency of Data.** The following recommendations pertain to changes in the peer-review
process that are aimed to improve the credibility and utility of DIAL data. Surprisingly few articles provided a citation to justify the examination of specific DIALs. We recommend that authors provide credible scientific or clinical arguments and supporting citations to justify the ethically significant decision to report data on some indicators of EoL care while omitting other indicators. We also recommend that journal editors consider the age of the data when making editorial decisions. The median timeliness index was about 3.5 years, and some studies reported data that were more than a decade old. Finally, we tried to identify researchers’ motives (e.g., show that palliative care is beneficial) but found that it was not possible to do so reliably. Future reports in this area would be more actionable if authors’ motives for conducting research were transparent.

Recommendations for Improving Research Implementation

In addition to offering recommendations for improving DIAL assessments, we offer some suggestions related to study design. Specifically, we call for more geographically dispersed prospective/TH research, more rigorous retrospective research, and greater use of blended designs.

Prospective and Terminal Hospitalization Designs. Prospective and terminal hospitalization designs are more flexible than the retrospective design, and permit more timely reporting of findings. The median timeliness index was greater than three years in retrospective studies (45 months) and less than three years in prospective studies (27 months). The interval between the last death observed in a cohort and the date of publication was shorter in prospective studies. This is important because changes in the quality of EoL care have been detected over intervals as brief as three years. Prospective and TH designs also permit the acquisition of data on variables that are rarely available in electronic health records or administrative data sets. Little is known about disparities in DIALs as a function of LGBTQ status, immigrant status, occupation, or language status. Data on these variables are rarely available in administrative data set or medical records but could be collected prospectively.

Prospective studies have typically been mounted in well-resourced settings by researchers affiliated with academic health centers (AHCs) who study patients receiving treatment in AHCs. These patients differ from those who do not seek treatment and those treated outside AHCs.

Going forward, prospective studies, including clinical trials, should be mounted in geographically dispersed regions, and they need to be much larger than they have been historically. Many of the apparently “null” effects in clinical trials are uninterpretable as these studies have lacked power to detect effects on utilization outcomes. Prospective and terminal hospitalization studies that cover patient and clinician populations that are large, geographically dispersed, and diverse will enable researchers to answer pressing questions that cannot be answered retrospectively. For example, use of those designs will enable scholars to discern whether differences in treatments within Dartmouth Atlas hospital referral regions are due to patient preferences, local norms of care, provider attitudes, or availability of resources (e.g., hospital beds). Given changes over time in demographic composition, and changes in societal, social, and family norms as well as norms of care, multilevel observational studies could be mounted at regular intervals by governments to eliminate disparities and improve EoL care.

Unlike Bach et al., we do not believe retrospective studies should be abandoned; rather, prospective and terminal hospitalization designs have methodological advantages that could inform the development of multilevel interventions. However, these designs require considerable resources and foresight, neither of which is essential to the conduct of credible research on DIALs. Retrospective studies are suitable alternatives when prospective research is not feasible economically or politically (e.g., low-resource settings, low- or middle-income countries).

Retrospective Designs. Beyond feasibility, retrospective studies have another considerable advantage. They were more likely to use government-sponsored administrative data sets that capture service use by people seen across disparate care settings, including community hospitals. Just as retrospective reconstructions of airplane crashes and other rare adverse events (e.g., suicide) have led to improvements in safety, retrospective research on DIALs can lead to improvements in EoL care.

If sampling is the main problem with the prospective design, a key problem with the retrospective design (e.g., using administrative data or medical charts) is measurement bias, especially (1) hindsight bias—the tendency to see something after-the-fact that is invisible prospectively; (2) rater bias—when one rater sees something that is invisible to other raters, due to their personal or professional history, perceptual acuity, or other characteristics; (3) information bias—when information that is available retrospectively or to one rater is unavailable prospectively or to another rater; and (4) artifacts attributable to differential item functioning—this can arise when independent raters use response categories in predictably different ways. For example, on a five-point scale assessing the avoidability of EoL treatments, ranging
from “highly likely” to “highly unlikely” with “somewhat likely,” “50/50” and “somewhat unlikely” in the middle, one raters’ “somewhat likely” will be mathematically equivalent to another rater’s “50/50” because they interpret the response categories differently.109

In one study, researchers minimized biases by having multiple raters from different medical specialties review medical records and answer open- and closed-ended questions about the avoidability of hospitalization. Critical care physicians were more likely than oncologists or hospitalists to view hospitalizations as avoidable, perhaps due to their work experiences. Alternatively, the apparent difference between specialists might be due to rater bias, information bias, or differential use of response scales. Biases can be accounted for statistically, by adjusting for rater characteristics.11 Artifacts can be minimized by design, using structured abstraction forms with evocative items or response options (e.g., anchoring vignettes).109

Blend the Designs. The advantages and disadvantages of prospective and retrospective designs in EoL research are well known.15,38—40 To increase the societal value of research,17 both designs ought to be evaluated by funders on how well they can attain multiple, often conflicting, deliverables. A major tradeoff is the prospective design’s potential to gain insight into the subjective experience of patients and family members who are at the EoL20 (often at great expense) versus the retrospective design’s capacity to yield larger, representative samples. Each design can include elements of the other, however. Prospective studies can include mortality follow-back components.149 Retrospective studies using electronic health records or administrative data sets can be linked to data sets collected prospectively using common data models such as PCORnet and/or assessments such as SEER-CAHPS,171 the Minimum Data Set,75 or PROMIS.172

Qualifiers
We note several qualifiers. First, the review covered previously unexplored terrain but was confined to English language articles, the unpublished, gray literature (e.g., technical reports, conference proceedings, dissertations) was not systematically reviewed, the protocol for this review was not published, and we have no quantitative data on relatively new treatments (e.g., targeted therapies). Future reviews can overcome all these problems. We make no claims about the review’s exhaustiveness; only studies reporting quantitative data on DIALs in the final 100 days of life were included, the exclusion of studies on children and persons with dementia is suboptimal, and we became aware of several eligible papers just as this manuscript was being completed.

Nonetheless, we believe that the inclusion of 254 articles covering research conducted in 29 countries and published in 79 journals over nearly 14 years is sufficiently comprehensive to justify our conclusions. Second, the statistical analyses did not account for the fact that some data sets, such as Medicare, yielded multiple publications. Third, we considered conducting a narrower review focused on one population (individuals with cancer) and a single methodology (analysis of administrative data sets). We opted for an expansive scoping review because there is a need for a less siloed discussion of methodological issues that cuts across multiple clinical specialties and research areas. Many dying patients suffer from multimorbidity, and EoL care needs to be improved across diagnostic groups at the population level, not just for patients with a particular disease.

Fourth, the task of compiling a list of potential indicators of DIALs is complicated by the fact that lists tend to ignore clinical context. For example, antibiotics may be discouraged in some dying patients but are also used for urinary tract infections and symptom management. Venting gastrostomies may be used for disease treatment or palliative purposes.174 In identifying potential DIALs in Appendix IV, we opted for a more inclusive approach, one that is commonly used in early stages of improving the measurement of latent constructs. Fifth, metrics that capture quality at the population level may not capture quality of care delivery for a particular patient at a specific time point. That tension should not prevent the development of population-based metrics, but it does underscore the limitations of the clinical utility (and credibility) of metrics. Sixth, it is likely that cognitively impaired patients and those who lacked decisional capacity at some point in their treatment were included in eligible studies. Finally, this review excluded patient-reported and caregiver-reported judgments of the quality of EoL care. Although these subjective judgments are important,175 we focused on widely available, universally applicable utilization outcomes because societies are accountable to all its citizens, including those who cannot speak for themselves and those who die without surviving caregivers or relatives.176

Conclusion
This scoping review identified more than 100 variables that have been studied over the final 100 days of life, prioritized important research areas, and analyzed critical conceptual and methodological issues. Operational definitions of DIALs will continue to evolve with changes in available treatments, health care delivery, and health policy. Our findings underscore the needs to engage the public and reflect
clinical realities when measuring DIALs for clinical, administrative, health policy, or research purposes.

**Disclosures and Acknowledgments**

The work reported here was partially supported by grants from the National Cancer Institute (R01CA168387, R01CA140419). The authors wish to acknowledge the contributions of librarian Karen Liljequist and coders Daniel Chess, Morgan McDonald, Spencer Obrecht, Meltem Otuncemur, Grant Reggio, Camila Rivero-Fernández, and Elyssa Steinberg. Michelle Ruidiaz-Santiago provided editorial assistance. Marsha Wittink, MD, MBe provided helpful input in the early stages of this project. Robert Horowitz, MD, and Susan Ladwig, MPH critiqued a draft of this manuscript.

The authors have nothing to disclose.

**References**


40. Bach PB, Schrag D, Begg CB. Resurrecting treatment histories of dead patients: a study design that should be laid to rest. JAMA 2004;292:2765–2770.


### Appendix I. Coding Form

<table>
<thead>
<tr>
<th>Var 0 Study Number</th>
<th>__ __ __ __ __ __ __ __ Whole number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Var 1 Journal Name</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>If other, write out:</td>
<td></td>
</tr>
<tr>
<td>Var 2 Professional Psych Society</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 3 # authors</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 4 Psychologist Author</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 5 Citation</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 6 Date of Last Death</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 7 Date of Publication</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 8 Number of deaths</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 9 Older adults only</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 10 Percent men (deaths only)</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 11 Diagnosis</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 12 Country</td>
<td>Separate coding sheet</td>
</tr>
<tr>
<td>Var 13 State</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __ See list of provinces</td>
</tr>
<tr>
<td>Var 14 Province</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 15 Design</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 16 Administrative Dataset</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 17 Medicare (US only)</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 18 Site of death</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 19 Inpatient admission</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var20 ED visits</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 21 ICU stay</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var22 Chemotherapy</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 23 CPR/resuscitation</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 24 Mechanical ventilation</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 25 Dialysis/hemodialysis</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 26 Intubation</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 27 Tracheotomy</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 28 Feeding tube</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 29 Hospice</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 30 Palliative care</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 31 Medication underuse/overuse</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>List medications studied (not cancer chemo):</td>
<td>Describe:</td>
</tr>
<tr>
<td>Var 32 Other—unknown</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 33 Other comfort care includes, but not limited to, care received from nurses, social workers, counselors, psychiatrists, psychologists, and clergy. If yes, describe:</td>
<td>Describe:</td>
</tr>
<tr>
<td>Var 34 Other invasive/life-prolonging measures</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>If yes, describe</td>
<td></td>
</tr>
<tr>
<td>Var 35 Chemotherapy ≤1 month before death</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 36 Chemotherapy ≤14 days before death</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 37 Number of days before death is ≤1</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 38 2 days before death</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 39 3 days before death</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 40 7 days before death</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 41 2 weeks before death</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 42 1 month (30–31 days) before death</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 43 2 months before death</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 44 More than 2 months before death</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 45 Terminal hospitalization</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Var 46 and Var 47</td>
<td>BLANK</td>
</tr>
<tr>
<td>Var 48 Effects of psychological variable</td>
<td>__ __ __ __ __ __ __ __ __ __ __ __ __</td>
</tr>
<tr>
<td>Describe:</td>
<td></td>
</tr>
</tbody>
</table>
Var 49 Effects of sex — 1 = yes, 2 = no
Var 50 Effects of age — 1 = yes, 2 = no
Var 51 Effects of socioeconomic status — 1 = yes, 2 = no
Var 52 Effects of race/ethnicity — 1 = yes, 2 = no
Var 53 Effects of religion/spirituality — 1 = yes, 2 = no
Var 54 Effects of medical comorbidity — 1 = yes, 2 = no
Var 55 Composite score — 1 = each indicator analyzed separately, 2 = created a composite or factor score 3 = combination of 1 and 2.

Countries

a) Australia 1 = yes, 2 = no
b) Belgium 1 = yes, 2 = no
c) Canada 1 = yes, 2 = no
d) China 1 = yes, 2 = no
e) Denmark 1 = yes, 2 = no
f) England/UK 1 = yes, 2 = no
g) Germany 1 = yes, 2 = no
h) India 1 = yes, 2 = no
i) Israel 1 = yes, 2 = no
j) Italy 1 = yes, 2 = no
k) Japan 1 = yes, 2 = no
l) Korea 1 = yes, 2 = no
m) Netherlands 1 = yes, 2 = no
n) New Zealand 1 = yes, 2 = no
o) Norway 1 = yes, 2 = no
p) Portugal 1 = yes, 2 = no
q) Singapore 1 = yes, 2 = no
r) Spain 1 = yes, 2 = no
s) Sweden 1 = yes, 2 = no
t) Taiwan 1 = yes, 2 = no—
u) United States 1 = yes, 2 = no
v) Other 1 = yes, 2 = no

Journals

a. BMJ/British Medical Journal
b. Cancer
c. Journal of the American Geriatrics Society
d. JAMA/Journal of the American Medical Association
e. JAMA Internal Medicine/Archives of Internal Medicine
f. Journal of Clinical Oncology
g. Journal of General Internal Medicine
h. Journal of Pain and Symptom Management
i. Journal of Palliative Care
j. Journal of Palliative Medicine
k. New England Journal of Medicine
l. Palliative Medicine
m. Other
n. BMC Palliative Care
o. BMJ Supportive and Palliative Care
p. Critical Care Medicine
q. Gynecologic Oncology
r. Journal of Oncology Practice
s. Medical Care
t. The Oncologist
u. Supportive Care in Cancer
Appendix II. Alphabetical List of Included Studies


Penders YW, van den Block L, Donker GA, Deliens L, Onwuteaka-Philipsen B, EURO IMPACT. Comparison of end-of-life care for older people living at home and in residential homes: a mortality follow-back study among GPs in The Netherlands. Br J Gen Pract 2015;65:e724-e730


Setoguchi S, Glynn RJ, Stedman M, Flavell CM, Levin R, Stevenson LW. Hospice, opiates, and acute care service use among the elderly before death from heart failure or cancer. Am Heart J 2010;160:139-144.


West E, Costantini M, Pasman HR, Onwuteaka-Philipsen B, EURO IMPACT. A comparison of drugs and procedures of care in the Italian hospice and hospital settings: the final three days of life for cancer patients. BMC Health Serv Res 2014;14:496.


Wong SP, Kreuter W, O'Hare AM. Treatment intensity at the end of life in older adults receiving long-term dialysis. Arch Intern Med 2012;172:661-663; discussion 663-664.


**Appendix III. Number (%) of Articles Reporting Data on DIALs Over Different EoL Timeframes**

<table>
<thead>
<tr>
<th>Timeframe</th>
<th>Retrospective</th>
<th>Prospective</th>
<th>Terminal Hospitalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final day of life</td>
<td>6 (3.6)</td>
<td>1 (2.2)</td>
<td>8 (17.0)</td>
</tr>
<tr>
<td>Final week of life</td>
<td>26 (15.6)</td>
<td>29 (63.0)</td>
<td>10 (21.3)</td>
</tr>
<tr>
<td>Final 14 days of life</td>
<td>53 (31.7)</td>
<td>13 (28.3)</td>
<td>8 (17.0)</td>
</tr>
<tr>
<td>Final month of life</td>
<td>113 (67.7)</td>
<td>15 (32.6)</td>
<td>9 (19.1)</td>
</tr>
</tbody>
</table>

EoL = end of life.

Bolded values in a column are significantly different from the values in the other two columns.

There were 167 retrospective cohorts, 46 prospective cohorts, and 47 terminal hospitalization cohorts.

*Most (n = 38, 80.9%) of these studies did not report a specific time interval before death.

$^1$Pearson chi² = 23.86, df = 1, $P < 0.001$.

$^2$Pearson chi² = 43.14, df = 1, $P < 0.001$.

$^3$Pearson chi² = 41.99, df = 1, $P < 0.00$.
### Appendix IV. Discretionary Interventions That Were Not Prespecified

<table>
<thead>
<tr>
<th>Category</th>
<th>Examples</th>
<th>Category</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>General support and monitoring</td>
<td>Arterial blood gas&lt;br&gt;Blood draw/blood test/blood work/lab tests&lt;br&gt;Catheter placement&lt;br&gt;Central line/vascular access placement&lt;br&gt;Imaging/radiologic testing/radiologic examination&lt;br&gt;IV fluids/IV hydration/IV infusion/microbiology&lt;br&gt;Minor respiratory therapeutic procedure&lt;br&gt;Secretion management&lt;br&gt;Urinary catheter/urinary catheterization&lt;br&gt;Vascular access placement/central line/peripherally Vital signs</td>
<td>Invasive diagnostics</td>
<td>Biopsy&lt;br&gt;Bronchoscopy/bronchus biopsy&lt;br&gt;Cardiac angiogram&lt;br&gt;Colonoscopy&lt;br&gt;Cystoscopy&lt;br&gt;Electromyelogram&lt;br&gt;Endoscopic retrograde cholangiopancreatography&lt;br&gt;Endoscopy/upper gastrointestinal endoscopy&lt;br&gt;Esophagogastroduodenoscopy&lt;br&gt;Laryngoscopy&lt;br&gt;Lumbar puncture/diagnostic spinal tap&lt;br&gt;Nephrostogram&lt;br&gt;Procedures requiring pathology consult&lt;br&gt;Transesophageal echocardiogram&lt;br&gt;Urinary tract—therapeutic procedures&lt;br&gt;Urogram</td>
</tr>
<tr>
<td>Higher intensity organ support and monitoring</td>
<td>Central venous catheter/venous access devices/central catheter&lt;br&gt;Central venous pressure measurement&lt;br&gt;Defibrillation/cardioversion&lt;br&gt;Hemodialysis catheter&lt;br&gt;Insertion of emergency airway&lt;br&gt;Intra-aortic balloon pump&lt;br&gt;Pulmonary artery catheter&lt;br&gt;Therapy/boluses&lt;br&gt;Ventricular assist device</td>
<td>Disease treatments</td>
<td>Chemoradiation&lt;br&gt;Hormonal therapy&lt;br&gt;Injection of granulocyte colony-stimulating factor&lt;br&gt;Radiation/radiotherapy&lt;br&gt;Surgery/surgical procedures&lt;br&gt;Transfusion of red blood cells/platelets/blood products</td>
</tr>
<tr>
<td>Fluid removal</td>
<td>Abdominal drainage/abdominal paracentesis/paracentesis&lt;br&gt;Chest drain/pleural drain/thoracentesis/pleural-peritoneal drain&lt;br&gt;Incision of pleura&lt;br&gt;Intrapertioneal drain&lt;br&gt;Pleurodesis&lt;br&gt;Pericardiocectomy&lt;br&gt;Incision of pleura</td>
<td>Managing obstructions</td>
<td>Cholecystectomy tube placement&lt;br&gt;Nephrostomy/nephrostomy tube placement&lt;br&gt;Obstructive uropathy&lt;br&gt;Percutaneous nephrostomy&lt;br&gt;Percutaneous transhepatic cholangiography and drainage&lt;br&gt;Removal of urinary obstruction&lt;br&gt;Small bowel resection&lt;br&gt;Transurethral excision&lt;br&gt;Urethral stent placement</td>
</tr>
<tr>
<td>Wound management</td>
<td>Debridement of wound, infection, burn, artery&lt;br&gt;Excision of skin lesion&lt;br&gt;Incision and drainage&lt;br&gt;Percutaneous enterogastric tube&lt;br&gt;Jejunostomy&lt;br&gt;Colorectal procedures/gastroenterology procedures&lt;br&gt;Gastrostomy/percutaneous gastrostomy/G-tube</td>
<td>Bleeding and clots</td>
<td>Ligation of vessel&lt;br&gt;Ablation therapy&lt;br&gt;Inferior vena cava filter placement&lt;br&gt;Injection of epinephrine</td>
</tr>
<tr>
<td>Gastrointestinal, gastrostomy and jejunostomy tubes*</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The words and phrases that appear as examples in this table were extracted from articles that were included in the review. In some cases, researchers merely reported descriptive data on EoL utilization without framing these activities as discretionary (or inappropriate). In the interest of completeness, we included all of these interventions, recognizing that many of them are symptom-focused activities that are considered vital to caring for patients who are dying.

*Can be from obstructive causes or other but not artificial nutrition.