

A Methodological Analysis of Chaplaincy Research: 2000–2009

KATHLEEN GALEK and KEVIN J. FLANNELLY

The Spears Research Institute, Healthcare Chaplaincy, New York, New York, USA

KATHERINE R. B. JANKOWSKI

Professional and Continuing Studies, Healthcare Chaplaincy, New York, New York, USA

GEORGE F. HANDZO

Chaplaincy Care Leadership & Practice, Healthcare Chaplaincy, New York, New York, USA

The present article presents a comprehensive review and analysis of quantitative research conducted in the United States on chaplaincy and closely related topics published between 2000 and 2009. A combined search strategy identified 49 quantitative studies in 13 journals. The analysis focuses on the methodological sophistication of the studies, compared to earlier research on chaplaincy and pastoral care. Cross-sectional surveys of convenience samples still dominate the field, but sample sizes have increased somewhat over the past three decades. Reporting of the validity and reliability of measures continues to be low, although reporting of response rates has improved. Improvements in the use of inferential statistics and statistical controls were also observed, compared to previous research. The authors conclude that more experimental research is needed on chaplaincy, along with an increased use of hypothesis testing, regardless of the research designs that are used.

KEYWORDS *chaplaincy, pastoral care, research methodology, quantitative research, statistical analyses*

Address correspondence to Kathleen Galek, Ph.D., The Spears Research Institute, Healthcare Chaplaincy, 307 East 60th Street, New York, NY 10022. E-mail: kgalek@healthcarechaplaincy.org

INTRODUCTION

As health care in the United States and elsewhere becomes more evidence-based, all clinical services, including chaplaincy, must conduct credible research to demonstrate their value and effectiveness. Although there is very little research on the effectiveness of professional chaplains, there is a small but growing body of research on how chaplains work with patients and families, and how satisfied patients and families are with the chaplain services.

Methodological reviews of the quality of research on chaplaincy, pastoral care, and related topics have been published in chaplain and pastoral care journals in the past. A review by Gartner, Larson, and Vachar-Mayberry (1990) evaluated the quality of quantitative research published between 1975 and 1984 in the the *Journal of Pastoral Care*, the *Journal of Pastoral Counseling*, *Pastoral Psychology*, and the *Journal of Religion and Health*. A subsequent review by Flannelly, Liu, Oppenheimer, Weaver, and Larson (2003) analyzed the methodological sophistication of quantitative research published between 1990 and 1999 in the the *Journal of Pastoral Care*, *Pastoral Psychology*, and the *Journal of Religion and Health*. The later review also examined changes in research quality over time. However, chaplaincy was not the sole focus of these reviews, which included research on other aspects of pastoral care, as well as research on the relationship between religion and health. A related review by Weaver, Flannelly, and Liu (2008) compared the methodological sophistication of this research to research on religion and health published between 1985 and 2002, in selected medical and social science journals. The interested reader is advised to consult a review by O'Connor et al. (2001) that critiqued qualitative research on pastoral care and counseling during the 1990s.

The present article presents a comprehensive review and analysis of quantitative research on chaplaincy conducted in the United States and published between 2000 and 2009. The analysis focuses on the methodological sophistication of this research and compares the quality of current (2000–2009) and past research.

METHODS

The literature review involved several search elements and levels of decision-making. First, every issue of *Chaplaincy Today*, *Pastoral Psychology* and the *Scottish Journal of Health Care Chaplaincy* was read because they are not indexed on Medline. Second, electronic searches were conducted on Medline for all the articles published in 2000–2009 in the *Journal of Health Care Chaplaincy*, the *Journal of Pastoral Care and Counseling*, and *Journal of Religion and Health*, using the search-phrase: (religion [major] OR spirituality [major]) AND (pastoral care OR chaplain*). Finally, an electronic

search of all articles indexed on Medline was conducted for the years 2000–2009, using the same search phrase. Both electronic searches were restricted to “journal articles” in English.

A decision tree was used for selecting articles for analysis. First, each article’s title was read to see if it was relevant. If the title clearly indicated the article was not about chaplaincy, the article was eliminated from further consideration. If this could not be determined, a researcher read the article’s abstract, and decided whether to include it or not. Based on the title and abstract, the researcher also decided whether the article was a quantitative study. If it was not, the article was excluded. Researchers then read all the quantitative studies, and all articles for which it was still uncertain whether (a) they were quantitative or (b) they were about chaplaincy. Finally, the researchers excluded any study that was not conducted in the United States.

The quantitative studies were classified into 11 categories: (1) attitudes about the roles of chaplains; (2) the nature of chaplain visits and interventions; (3) referrals to chaplains; (4) instrument development; (5) patient and family satisfaction with chaplains; (6) interventions studies; (7) other studies with patient populations; (8) chaplaincy staffing; (9) chaplain attitudes and perceptions; (10) chaplain well-being; and (11) chaplain education.

RESULTS

The literature review identified 49 quantitative studies in 13 journals. The *Journal of Pastoral Care and Counseling* (JPCC) published more than half of the studies in the sample (51.0%), and the *Journal of Health Care Chaplaincy* (JHCC) (16.3%) and *Chaplaincy Today* (CT) (12.2%) published over a quarter of them. The remaining 10 journals each published one quantitative chaplaincy study: *Holistic Nursing Practice* (HBP), *Hospital Topics* (HT), the *Journal of Nervous & Mental Disease* (JNMD), the *Journal of Religion and Health* (JRH), *Mayo Clinic Proceedings* (MCP), *Mental Health, Religion & Culture* (MRC), *Psycho-Oncology* (PO), *Pastoral Psychology* (PP), the *Scottish Journal of Health Care Chaplaincy* (SJHCC), and the *Southern Medical Journal* (SMJ).

Descriptions of Studies

Table 1 lists the studies conducted on chaplain roles, chaplain visits and interventions, and referrals to chaplains. Many of the articles in Table 1 are described in greater detail in Jankowski et al. (2008).

ATTITUDES ABOUT CHAPLAIN ROLES

Two national surveys of the attitudes of hospital administrators about the roles of chaplains were conducted early in the decade (Table 1). The first,

TABLE 1 Chaplain Roles, Activities, and Referrals

Year	Journal	First author	Data collection	Sampling method	Sample
<i>Attitudes about Chaplains' Roles</i>					
2005	JPCC	Flannelly	Survey	Random	Hospital Administrators
2005	HT	Flannelly	Survey	Random	Hospital Directors
2006	JPCC	Flannelly	Survey	Random	Hospital Directors
<i>Chaplaincy Visits and Interventions</i>					
2003	PO	Flannelly	Records	Convenience	Chaplains
2004	CT	Flannelly	Records	Convenience	Chaplains
2005	CT	Spidell	Survey	Convenience	Chaplains
2008	JHCC	Hummel	Survey	Convenience	Chaplains
2008	JHCC	Handzo	Records	Convenience	Chaplains
2008	JHCC	Handzo	Records	Convenience	Chaplains
2008	JHCC	Vanderwerker	Records	Convenience	Chaplains
<i>Referrals to Chaplains</i>					
2000	JPCC	Fitchett	Survey	Convenience	In-Patient
2004	JPCC	Fogg	Records	Convenience	Chaplains
2007	MHRC	Galek	Survey	Random	Hospital Directors
2008	JHCC	Vanderwerker	Records	Convenience	Chaplains
2009	JPCC	Galek	Records	Convenience	Chaplains

published in 2005, asked U.S. hospital executives and pastoral care directors about the importance of different chaplain roles and functions (Flannelly, Handzo, Weaver, & Smith, 2005). The next two reports listed in Table 1 present findings about chaplain roles based on a survey of hospital directors of medicine, nursing, social work and pastoral care (Flannelly, Galek, Bucchino, Handzo, & Tannenbaum, 2005; Flannelly, Handzo, Galek, Weaver, & Overvold, 2006). The survey questionnaires were developed from previous research.

CHAPLAIN VISITS AND INTERVENTIONS

Seven studies were found that described chaplain visits and interventions (see Table 1). The first two articles in this category (Flannelly, Weaver, & Handzo, 2003, 2004) are an original report and a reprinted report of the findings from a New York City hospital that was part of a larger study conducted in the late 1990s: The New York Chaplaincy Study (NYCS). The study analyzed computerized records describing health status, religion, and other patient variables and chaplain interventions. The two 2008 studies by Handzo and his colleagues report the findings from the entire study, which recorded over 30,000 patient visits by dozens of chaplains and clinical pastoral education (CPE) students at a dozen hospitals in the New York City area (Handzo, Flannelly, Kudler, et al., 2008; Handzo, Flannelly, Murphy, et al., 2008). The final study in this section of the table compared selected findings from the NYCS and the Manhattan Chaplaincy Study (MCS), which was conducted in 2005–2008 and included close to 120,000 patient visits. The results presented in this article by Vanderwerker and her colleagues (2008)

were based on over 33,000 visits from the NYCS and nearly 60,000 chaplain visits from the MCS.

The remaining two studies in the second section of Table 1 were surveys of small samples of chaplains. The first asked chaplains about their beliefs and practices (Spidell, 2005), while the second asked chaplains to rate the degree to which they thought a list of nursing interventions constituted spiritual care (Hummel, Galek, Murphy, Tannenbaum, & Flannelly, 2008).

The studies based on hospital records are observational studies in a sense, since the chaplains and CPE students involved in the data collection observed and recorded their own visits and activities with patients, families, friends, and staff. The chaplains also collected data about the patients they visited, which were analyzed in the studies, but the unit of analysis was visits.

REFERRALS TO CHAPLAINS

The bottom section of Table 1 contains five studies that examined referrals to chaplains. Two of them were surveys (Fitchett, Meyer, & Burton, 2000; Galek, Flannelly, Koenig, & Fogg, 2007) and three of them analyzed computerized records of referrals that were reported by chaplains and CPE students (Fogg, Weaver, Flannelly, & Handzo, 2004; Galek et al., 2009; Vanderwerker et al., 2008). One of the surveys asked random samples of department directors of nursing, medicine, social work and pastoral care how important they thought it was to refer patients to chaplains for different kinds of issues (Galek et al., 2007). The other asked patients if they wanted different kinds of chaplaincy services (Fitchett et al., 2000).

The sample sizes of the survey studies were roughly 200 and 1200 persons, respectively. The unit of analysis in the Fogg et al. (2004), Vanderwerker et al. (2008), and Galek et al. (2009) studies was chaplain visits with patients or family members that were made in response to referrals. The number of referred visits was roughly 3,700, 7,000, and 15,600 in the three studies.

INSTRUMENT DEVELOPMENT

The Medline search identified several articles on the development and testing of scales to measure religion and spirituality, which might be valuable, but were not deemed to be directly relevant to the current analysis. Four other studies, which developed and tested scales, were deemed to be relevant (see Table 2). The studies examined three different types of scales: spiritual needs (Flannelly, Galek, & Flannelly, 2006; Grosseohme, 2008), spiritual struggle (Fitchett & Risk, 2009), and family satisfaction with chaplains (Flannelly, Galek, Tannenbaum, & Handzo, 2007). The instruments developed by Grosseohme (2008) and Fitchett and his colleagues (2009) were specifically designed as screening tools to determine if patients need

TABLE 2 Measurement Instruments for Chaplaincy

Year	Journal	First author	Data collection	Sampling method	Sample
2006	HNP	Flannelly	Survey	Convenience	Chaplains
2007	JPCC	Flannelly	Survey	Convenience	Chaplains
2008	JPCC	Grossoehme	Survey	Convenience	Chaplains/Children*
2009	JPCC	Fitchett	Survey	Convenience	In-Patients

*Chronically ill hospitalized children.

chaplaincy care. The pilot testing of the family satisfaction scale developed by Flannelly et al. study (2007) and the spiritual needs scale for children developed by Grossoehme (2008) were very limited.

PATIENT AND FAMILY SATISFACTION

Four studies on patient and family satisfaction with pastoral care in hospital settings are listed in Table 3. Several of the articles in Table 3 are described in greater detail by Jankowski, Handzo, and Flannelly (2011).

The first article presents the results of a large survey of patient satisfaction with chaplaincy services at several U.S. hospitals (VandeCreek, 2004). The study, which surveyed over 1,400 patients, represents the culmination of a series of studies conducted in the 1990s to develop a patient satisfaction questionnaire. The second study asked 535 patients about their expectations of, and satisfaction with, chaplaincy services at the Mayo Clinic (Piderman et al., 2008). The third study measured the extent to which patients felt a chaplain visit was effective in meeting their spiritual and emotional needs of 200 patients (Flannelly, Oettinger, Galek, Braun-Storck, & Kreger, 2009). The last study measured the satisfaction of 130 bereaved family members with the care hospital chaplains provided to them and their deceased loved one (Broccolo & VandeCreek, 2004).

TABLE 3 Patient and Family Satisfaction and Intervention Studies

Year	Journal	First author	Data collection	Sampling method	Sample
<i>Satisfaction Studies</i>					
2004	JPCC	VandeCreek	Survey	Convenience	Discharged Patients
2008	MCP	Piderman	Survey	Convenience	In-Patients
2009	JPCC	Flannelly	Survey	Convenience	In-Patients
2004	JPCC	Broccolo	Survey	Convenience	Family Members
<i>Intervention Studies</i>					
2000	JPCC	Tartaglia	Quasi-Expmt	Convenience	Adults
2006	JPCC	Tibbits	Experiment	Convenience	Adults
2008	JPCC	Farrell	Experiment	Convenience	Children
2001	CT	Iler	Experiment	Convenience	Adults
2008	JRH	Bay	Experiment	Random	Adults

The study by Flannelly et al. (2007) adapted several items from the VandeCreek (2004) patient satisfaction scale, in addition to utilizing a number of untested items. Piderman and his colleagues (2008) developed their own questionnaire items, as did Broccolo and VandeCreek (2004).

INTERVENTION STUDIES

Five studies that reported the findings of various interventions with patients or family members are also listed in Table 3. The first study in this section of the table examined a pastoral care intervention to increase organ donations (Tartaglia & Linyear, 2000), and the second tested a chaplain-run forgiveness training program to reduce hypertension in 25 patients (Tibbits, Ellis, Piramelli, Luskin, & Lukman, 2006). The third study used a religiously based play intervention to reduce anxiety and depression among a sample of forty chronically ill hospitalized children (Farrell, Cope, Cooper, & Mathias, 2008).

The remaining two articles present the results of the few experimental studies to measure the effectiveness of chaplain interventions with patients. An early experiment by Florell (1973) appears to be the only other experimental study on the effectiveness of a chaplains working with patients, but it did not use a chaplain intervention, per se (Jankowski et al., 2011).

The 2001 study by Iler, Obenshain, and Camac found that a chaplain visit of approximately 20 minutes significantly decreased anxiety and hospital length of stay, and increased satisfaction with care among 25 experimental (visited) patients compared to 25 control (non-visited) patients. Patients in the experimental group of the study by Bay, Beckman, Trippi, Gunderman, and Terry (2008) received five standardized chaplain visits of almost 45 minutes each (on average) over the course of several days. The study found no significant effects of chaplain visits on anxiety, depression or length of stay. However, chaplain visits were reported to enhance positive religious coping.

OTHER STUDIES WITH PATIENT POPULATIONS

Table 4 shows the remaining studies that have been conducted with patient populations that were published between 2000 and 2009. The three studies

TABLE 4 Other Studies of Patient Populations

Year	Journal	First author	Data collection	Sampling method	Sample
2001	JPCC	Grossoehme	Survey	Convenience	Adolescent Psychiatric Patients
2003	CT	Blanchard	Survey	Convenience	Cardiac Patients
2006	JPCC	Levy	Survey	Convenience	Cancer Patients
2007	JPCC	Grossoehme	Survey	Convenience	Adolescent. In-Patients
2008	JHCC	Grossoehme	Survey	Convenience	Adolescent Out-Patients

by Grossoehme and his colleagues surveyed adolescent patients about their spirituality and one compared the spirituality of hospitalized adolescents to that of other adolescents (Grossoehme, 2001; Grossoehme, Cotton, & Leonard, 2007; Grossoehme, VanDyke, & Seid, 2008). The sample size of the three studies ranged between approximately 120 and 200 patients. Blanchard (2003) assessed the spiritual needs of 161 cardiac patients, and Levy and Chan (2006) studied spiritual coping in 13 cancer patients.

CHAPLAINCY STAFFING

Table 5 includes several national surveys that have looked at the staffing of chaplaincy departments and the factors that affect it. Although the samples of these studies on *Staffing* are listed as Pastoral Care Directors and Hospital Administrators, the administrators were providing information about their hospitals, and it was the information about the hospitals that was the focus of each study.

The study by VandeCreek (2000) asked close to 400 pastoral care directors about their staffing levels, budgets, and facilities. A followup report (VandeCreek, Siegel, Gorey, Brown, & Toperzer, 2001) calculated the number of chaplains per patient and explored some of the variables that were associated with staffing levels. A 2004 survey of nearly 500 pastoral care directors (Flannelly, Handzo, & Weaver, 2004) replicated and extended the work of VandeCreek, Siegel, Gorey, Brown and Toperzer (2001), using regression analyses to estimate the relative contributions of hospital size, location, religious affiliation, profit/non-profit status, and other factors on the number and type of staff (board certified) per patient. Cadge and her colleagues further replicated and extended Flannelly et al.'s (2004) analyses, using data from the American Hospital Association's annual survey, to assess the degree to which a similar set of variables predicted whether or not a hospital had a pastoral care department (Cadge, Freese, & Christakis, 2008).

TABLE 5 Chaplaincy Staffing, Attitudes and Perceptions

Year	Journal	First author	Data collection	Sampling method	Sample
<i>Staffing</i>					
2000	JHCC	VandeCreek	Survey	Random	Pastoral Care Directors
2001	JPCC	VandeCreek	Survey	Random	Pastoral Care Directors
2004	JPCC	Flannelly	Survey	Random	Pastoral Care Directors
2008	SMJ	Cadge	Survey	Random	Hospital Administrators
<i>Attitudes and Perceptions</i>					
2003	CT	Pfannenstiel	Survey	Convenience	Chaplains
2005	JPCC	Dodd-McCue	Survey	Convenience	Chaplains
2006	SJHCC	Flannelly	Survey	Convenience	Chaplains

CHAPLAIN ATTITUDES AND PERCEPTIONS

Research measuring chaplains' attitudes and perceptions, apart from those about chaplain roles and functions, are also listed on Table 5. The first study in this section of the table asked 122 chaplains about their use of information resources (Pfannenstiel, 2003), and the second asked seven chaplains about their roles as members of an organ donation program (Dodd-McCue & Tartaglia, 2005). The last study asked 167 chaplains how often they encountered patients with various kinds of spiritual needs (Flannelly, Galek, Bucchino, & Vane, 2006). The Flannelly, Galek et al. (2006) study used the spiritual needs scale that was mentioned in the section on instrument development (Flannelly, Galek, et al., 2006; Galek, Flannelly, Vane, & Galek, 2005).

CHAPLAIN WELL-BEING

Table 6 includes an assortment of survey studies about education, and the emotional and spiritual well-being of chaplains. The first study on *Well-Being* explored the factors that contribute to the professional satisfaction of 1,000 members of the Association of Professional Chaplains. The next two articles report the findings of a survey about burn-out and secondary post-traumatic stress among 350 chaplains and other clergy one year after the 911 attacks (Flannelly, Roberts, & Weaver, 2005; Roberts, Flannelly, Weaver, & Figley, 2003). The study by Taylor and her colleagues measured these same variables in a separate sample of 66 Jewish chaplains (Taylor, Flannelly, Weaver, & Zucker, 2006). The last study asked 188 male and female chaplains about their spiritual needs in times of distress (Galek, Flannelly, Jacobs, & Barone, 2008).

CHAPLAIN EDUCATION

The first study in the *Education* section of Table 6 surveyed nearly 400 Canadians about their experiences as supervisory students (Trothen, 2001).

TABLE 6 Chaplain Spiritual and Emotional Well-Being and Education

Year	Journal	First author	Data collection	Sampling method	Sample
<i>Well-Being</i>					
2002	JPCC	Crossley	Survey	Convenience	Chaplains
2003	JNMD	Roberts	Survey	Convenience	Chaplains and Clergy
2005	JPCC	Flannelly	Survey	Convenience	Chaplains and Clergy
2006	JPCC	Taylor	Survey	Convenience	Chaplains
2008	JPCC	Galek	Survey	Convenience	Chaplains
<i>Education</i>					
2001	JPCC	Trothen	Survey	Convenience	Supervisory CPE Students
2004	CT	Grossoehme	Records	Convenience	Candidates for APC BC
2005	PP	Moran	Survey	Convenience	Clergy
2008	JHCC	Jankowski	Survey	Convenience	CPE students

The second examined APC records of 75 chaplains to determine the most common unmet competencies of first-time candidates for board certification (Grossoehme, 2004). The third study, which surveyed approximately 180 clergy about their pastoral care practices, found that clergy who had clinical pastoral education (CPE) felt more competent to provide pastoral counseling to their congregants for various kinds of problems (Moran et al., 2005). The final article is a quasi-experimental study of 140 CPE students that assessed changes in self-reflection, emotional intelligence and pastoral care skills associated with one unit of CPE training (Jankowski, Vanderwerker, Murphy, Montonye, & Ross, 2008).

Analysis of Studies

In all, the sample consisted of 36 surveys, eight observational studies, four experimental studies, and one quasi-experimental study. Table 7 shows how the studies fared in terms of their research sophistication. Eight quality indicators are listed in Table 7, some of which are more stringent than others. All eight have been used, to one degree or another, in previous assessments of research quality in chaplaincy, pastoral care, and religion and health (Flannelly, Liu et al., 2003; Gartner et al., 1990; Weaver, Flannelly, Strock, Krause, & Flannelly, 2005; Weaver et al., 2008).

SAMPLES

Twenty three of the 49 studies were about patients (46.9%), including the eight observational studies in which chaplains collected the data on hospital forms (Data Collection = Records in Table 1). The vast majority of all the studies used convenience samples (85.7%), with only seven of the 36 surveys using random samples (see Table 7). Keep in mind that random sampling (i.e., randomly selecting potential participants) is not the same as randomly assigning participants to control and experimental groups, which is one of the hallmarks of a true experiment (Campbell & Stanley, 1963).

TABLE 7 Percent of Studies Exhibiting Research Sophistication According to Various Indicators

Indicators	Experimental [†]	Observational	Survey
Used Random Sampling	0.0	0.0	19.4
Reported Return/Response Rate	60.0	0.0	77.8
Reported Validity of Measures	80.0	0.0	11.1
Reported Reliability of Measures	20.0	0.0	33.3
Tested Specific Hypotheses	60.0	0.0	25.0
Offered an Explanatory Model	0.0	12.5	11.1
Used Inferential Statistics	80.8	87.5	69.4
Used Statistical Controls	20.0	0.0	33.3

[†]Includes one quasi-experimental study.

The sample sizes of the survey research ranged from 7 to 1,514, with the median being 188 participants. The sample sizes for the quasi-experimental and experimental studies were between 25 and 204, with the median being 50. The majority of experiments and more than three-quarters of the surveys reported the percent of people invited to participate in the study that actually did so (i.e., the participation rate or return rate [see Table 7]). It is important to report response rates in order to be able to assess the likelihood of selection bias in a sample. We made no attempt to make such an assessment, however.

MEASURES

The studies used many different measures as dependent (i.e., outcome) variables, but few of the studies discussed the validity and reliability of the measures. Validity reflects the degree to which an instrument measures what is supposed to measure, and reliability reflects how consistent the measure is in doing so (Anastasi, 1982). To the extent that one employs commonly used and well-tested measures, this lack of reporting might not be considered a problem, but the authors of many studies developed their own measures.

No measures of validity or reliability were provided for any of the observational studies. Since the observational studies used measures that were created by chaplains to record chaplain activities, one would expect that they would at least have face validity. The fact that none of them discussed the reliability of their measures is more problematic, as inter-rater reliability is very important in such studies.

The hallmarks of a true experiment, according to Campbell and Stanley (1963), include (a) random assignment of individuals to different groups, (b) control of a manipulated intervention (or independent variable), and (c) measurement of the outcome (or dependent variable) with reliable instruments. The four true experiments in the current sample used validated scales as their dependent variables and discussed their validity. Only one, however, discussed the reliability of the scales they used. This experiment reported the Cronbach alpha's for the scales it used, which is a reliability measure of the internal consistency of a scale.

Hypotheses

Although virtually all of the studies stated the research question the study was intended to answer, few of them stated specific hypotheses (see Table 7). Specifying hypotheses was most common among the quasi-experimental and experimental studies, with three of the four experiments stating one or more specific hypotheses. Hypothesis testing was far less common in the survey research and none of the observational studies tested specific hypotheses. It was even less common for research papers in the sample to present explanatory models or theories that might account for their findings.

Statistics

Table 7 also shows the percent of studies that employed inferential statistics and statistical controls in their analyses. Inferential statistics enable one to know if the values observed for variables in a study differ greatly from what one would expect by chance. Statistical control means that extraneous variables that might influence an outcome variable are included in statistical analyses that are designed to measure the effects of, or association of, a specific variable with an outcome.

Between roughly 70% and 88% of the studies used inferential statistics, with the survey studies being the least likely to use them. Though the survey research was least likely to use inferential statistics, nearly half of those that did so used statistically controlled for one or more variables. Only one of the experiments and none of the observational studies statistically controlled for extraneous variables. The one experiment that used statistical controls was a study on the effectiveness of a chaplain intervention with patients (Bay et al., 2008), which controlled for patient demographic characteristics.

Table 8 show the different kinds of statistical tests used in the experimental, observational, and survey studies in the sample. Some of the studies used more than one kind of statistical test.

Roughly one-fifth of the surveys (19.4%) used relatively sophisticated statistical analyzes such as Analysis of Variance or Multivariate Analysis of Variance (Tabachnick & Fidell, 1996). Slightly over one-third used more sophisticated statistic methods, such as ANCOVA or MANCOVA (14%) or multiple regression (19.4%), which are specifically designed to control for extraneous variables (Tabachnick & Fidell). Their level of sophistication depends on the number of extraneous variables they actually include as control variables in the analyses.

TABLE 8 Percent of Studies using Various Types of Statistical Analyses[‡]

	Experimental [†]	Observational	Survey
ANOVA or MANOVA	20.0%	12.5%	19.4%
ANCOVA or MANCOVA	20.0%		13.9%
Chi-Square Test or Fisher's Exact Probability		62.5%	16.7%
Pearson's correlation coefficient (<i>r</i>)	20.0%	25.0%	8.3%
Student's <i>t</i> , Mann-Whitney <i>U</i> -test, Wilcoxon	80.0%		8.3%
Multiple Regression			19.4%
Exploratory Factor Analysis			13.9%

[†]Includes one quasi-experimental study.

[‡]Percentages exceed 100% because some studies performed more than one type of statistical analysis.

ANOVA = Analysis of Variance; ANCOVA = Analysis of Covariance; MANOVA = Multivariate Analysis of Variance; MANCOVA = Multivariate Analysis of Covariance.

Six survey studies used ordinary least squares regression and one used logistic regression. Chi-square and similar tests for nominal data were the third most widely used statistics for analyzing survey data.

The experimental studies mainly relied on different types of analysis of various designs (ANCOVA, MANOVA, etc.), which are commonly used in experimental research. Student's *t*-test, the Mann-Whitney *U*-test, and the Wilcoxon test were used for pair-wise comparisons in different experiments.

SUMMARY AND DISCUSSION

The foregoing analysis presents an overview of the current level of the methodological sophistication of chaplaincy research that can be at least partially compared to prior research in the field (Flannelly, Liu, et al., 2003; Gartner et al., 1990), as well as the methodological sophistication of research in selected related fields (Weaver et al., 2005). Although there is increasing interest in qualitative methods in chaplaincy (Fitchett, 2011), we focused on quantitative research because it remains the major methodology employed in chaplaincy and the health sciences.

Cross-sectional survey research continues to be the primary design employed in chaplaincy and pastoral care (Flannelly, Liu, et al., 2003; Gartner et al., 1990). However, surveys represented an even higher percent of studies on spirituality and health in 1985–2002 (Weaver et al., 2005). Experimental and quasi-experimental studies on chaplaincy are as rare as they were the 1990s (Flannelly, Liu, et al., 2003).

Campbell (1957) introduced the concept of external validity to determine whether the results of experimental and quasi-experimental research are generalizable. The concept has been extended to surveys and other kinds of studies (Flannelly, Liu, et al., 2003; Gartner et al., 1990; Weaver et al., 2008). External validity mainly hinges on the extent to which the sample under study is representative of the population to which one wants to generalize. A random sample is considered more likely to be representative of the population from which it is drawn than a convenience sample. Similarly, the higher the percentage of people who agree to participate in a study (i.e., the response rate), the more likely it is that the sample is representative. Finally, the larger its sample size, the more likely it is that the study's results will accurately reflect the population.

Changes in the research sophistication with respect to sampling have been mixed over the past 3–4 decades. Roughly 76% of all studies used convenience samples in the 1990s (Flannelly, Liu, et al., 2003) compared to 86% in the current sample. However, less than 20% of the pastoral care studies published in 1975–1984 even reported their sampling method (Gartner et al., 1990). The use of random samples in surveys was somewhat higher

in 1990–1999 (24%) than 2000–2009 (19%), but the median sample size of surveys has increased over time: 50 in 1975–1984; 150 in 1990–1999, and 186 in 2000–2009. Both the percent of surveys that employed random samples and their sample sizes were larger in research on spirituality and health examined by Weaver et al. (2005). Overall, improvements in sampling methods in chaplaincy research have been relatively minor and the field lags behind other fields of research in this respect.

With regard to response rates, the only information that has been examined from pastoral care and chaplaincy studies over the years (1975–2009) has been whether a study reported the response rate or not. Data on the response rates, or return rates, themselves, have not been collected (Flannelly, Liu, et al., 2003; Gartner et al., 1990). Nevertheless, there has been substantial progress made in reporting response rates. In 1975–1984 approximately one-third of studies reported response rates (Gartner et al., 1990), whereas the close to 43% of studies reported response rates in 1990–1999 (Flannelly, Liu, et al., 2003). This percentage rose to 63% in the current sample. Differences in reporting response rates were especially pronounced for survey studies in the 1999–1999 and 2000–2009 samples. While just over half of the surveys published in 1990–1999 reported response rates, more than three-quarters of the 2000–2009 surveys did so. However, since low response rates may be associated with response bias, which that can hamper generalization, greater attention must be paid to the actual response rates achieved in chaplaincy research.

The validity and reliability indicate the degree to which an instrument is measuring what it claims to be measuring and the accuracy with which it is doing so. As was the case with response rate, the data on validity and reliability used in this and previous analyses have been limited to whether these concepts are mentioned in the studies, not the extent to which the instruments meet standards of validity and reliability. Nevertheless, discussion of these concepts provides a measure of research sophistication, as such discussion typically gives some information about the degree to which the measures meet accepted standards.

Unfortunately, validity was rarely mentioned, except in the experimental studies, and reliability was mentioned relatively rarely among all the studies. Indeed the percentage of surveys in the current sample (33%) that mentioned reliability was no higher than that reported by Gartner et al. (1990) and Flannelly, Liu, et al. (2003) for their 1975–1984 and 1990–1999 samples.

The sophistication of the statistical analyses has also remained relatively unchanged since the decade of the 1990s. For example, 74% of all studies in the present sample used inferential statistics compared to 71% in 1990–1999. Among the survey studies, 69% published in 2000–2009 used inferential statistics compared to 77% published in 1990–1999. The percentages are substantially lower than those reported by Weaver et al. (2005) for research

on spirituality and health in 1985–1990 (91%) and 1997–2002 (100%), respectively.

The use of statistical controls is another important criterion for assessing the research sophistication of quantitative studies because the inclusion of multiple variables (such as demographic characteristics and other variables) in the statistical analyses makes it possible to determine the relative contribution of different variables on the outcomes being studied. Although the survey studies in the current sample were least likely to use inferential statistics, nearly half of those that did so statistically controlled for one or more variables. This is much higher than the 19% of surveys that controlled for one or more variables in their statistical analyses in the 1990s (Flannelly, Liu, et al., 2003), but it is much lower than the percent of survey studies in related fields that use statistical controls (Weaver et al., 2005).

Hypothesis testing and the use of explanatory models were relatively uncommon in the present sample. Only a few experimental studies tested specific hypotheses and only a quarter of the surveys did so. These findings are consistent with the low rate of hypothesis testing that has been observed in previous reviews of chaplaincy and pastoral care research (Flannelly, Liu, et al., 2003; Gartner et al., 1990). These percentages stand in marked contrast to the more frequent use of hypothesis testing in research on religion and health (Weaver et al., 2005).

Previous reviews of chaplaincy research have not looked at the use of explanatory models. However, Weaver and his colleagues (2005) found the use of explanatory models to help interpret the results was also infrequent in research on spirituality and health. As the field of chaplaincy, like the broader field of spirituality and health, lack theoretical frameworks from which explanatory models can be developed, it is not surprising that so few studies offer models to help explain their results. As such, the use of explanatory models may be a better measure of theory sophistication than of research sophistication per se. Hypothesis testing, on the other hand, is a critical measure of research sophistication, because testing the falsifiability of hypotheses is the only accepted standard of scientific progress (Popper, 1989, 1992).

CONCLUSIONS

Our analysis attempts, in part, to compare changes in the methodological sophistication of chaplaincy research across time. However, since previous reviews of the field have included studies that lie outside of chaplaincy, per se (Flannelly, Liu, et al., 2003; Gartner et al., 1990), these comparisons only provide approximations about the size of these changes. Nevertheless, we believe these comparisons provide useful insights about the

methodological progress of the field, and the progress that still has to be achieved.

Cross-sectional surveys have been and continue to be the mainstays of chaplaincy research (Flannelly, Liu, et al., 2003; Gartner et al., 1990). However, experimental and quasi-experimental studies are needed to adequately assess the efficacy of chaplain interventions, because causal inferences cannot be drawn from cross-sectional studies.

Three aspects of sampling are important for making generalizations from studies: (1) the type of sample (random sample versus other samples), (2) sample size, and (3) response rate. The vast majority of chaplaincy studies still employ convenience samples, which limits the generalizability of results. Although experimental studies randomly assign participants to conditions, the use of convenience samples still limits generalizability.

At least among survey studies, the median sample size of chaplaincy and pastoral care studies has gradually increased over time, which enhances the ability to generalize their results. However, median sample sizes still remain relatively small, and the samples sizes of surveys vary widely. Both the percent of surveys that employ random samples and their sample sizes need to increase to the levels observed among studies on spirituality and health (Weaver et al., 2005). While substantial progress has been made in reporting response rates, since low response rates may be associated with response bias that can hamper generalization to the larger population from which a sample was drawn, greater attention has to be paid to the actual response rates achieved in chaplaincy studies. The findings also indicate that chaplaincy research lags behind other fields with respect to reporting validity and reliability.

Improvements in the methodological sophistication with respect to statistical analyses have been mixed. The use of inferential statistics remained relatively unchanged between the 1990s and the 2000s, although surveys in the present sample were much more likely to control for one or more variables than those published in the 1990s (Flannelly, Liu, et al., 2003). Even so, chaplaincy surveys were less likely to employ statistical controls than related research on spirituality and health (Weaver et al., 2005). More sophisticated statistical analyses are necessary in chaplaincy research in order to be able understand the interplay of different variables and their relative contributions to health outcomes.

Hypothesis testing and the use of explanatory models continue to be relatively uncommon in research related to chaplaincy. It is understandable that explanatory models are rare, since there is very little theory from which to draw from in the field of chaplaincy. One might attribute the rarity of hypothesis testing to lack of theories in the field, as well. However, hypotheses need not be drawn from theories. They can be empirical, theoretical, or simply hunches. We urge researchers to increase their use of hypothesis testing, regardless of the research designs that they use.

REFERENCES

- Anastasi, A. (1982). *Psychological testing*. London, UK: Collier MacMillan.
- Bay, P. S., Beckman, D., Trippi, J., Gunderman, R., & Terry, C. (2008). The effect of pastoral care services on anxiety, depression, hope, religious coping, and religious problem solving styles: a randomized controlled study. *Journal of Religion and Health, 47*(1), 57–69.
- Blanchard, J. H. (2003). Covenanted companionship: Assessing spiritual needs with the open heart patient. *Chaplaincy Today, 19*(2), 10–15.
- Broccolo, G. T., & VandeCreek, L. (2004). How are health care chaplains helpful to bereaved family members? Telephone survey results. *Journal of Pastoral Care and Counseling, 58*(1–2), 31–39.
- Cadge, W., Freese, J., & Christakis, N. A. (2008). The provision of hospital chaplaincy in the United States: A national overview. *Southern Medical Journal, 101*(6), 626–630.
- Campbell, D. O. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin, 54*(4), 297–312.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally.
- Crossley, K. (2002). Professional satisfaction among U.S. healthcare chaplains. *Journal of Pastoral Care and Counseling, 56*(1), 21–27.
- Dodd-McCue, D., & Tartaglia, A. (2005). The impact of the Family Communication Coordinator (FCC) Protocol on the role stress of hospital chaplains. *Journal of Pastoral Care and Counseling, 59*(4), 345–360.
- Farrell, J., Cope, S. B., Cooper, J. H., & Mathias, L. (2008). Godly play: An intervention for improving physical, emotional, and spiritual responses of chronically ill hospitalized children. *Journal of Pastoral Care and Counseling, 62*(3), 261–271.
- Fitchett, G. (2011). Making our case(s). *Journal of Health Care Chaplaincy, 17*(1), 3–18.
- Fitchett, G., Meyer, P. M., & Burton, L. A. (2000). Spiritual care in the hospital: Who requests it? Who needs it? *Journal of Pastoral Care, 54*(2), 173–186.
- Fitchett, G., & Risk, J. L. (2009). Screening for spiritual struggle. *Journal of Pastoral Care and Counseling, 63*(1–2), 4–1–12.
- Flannelly, K., Liu, C., Oppenheimer, J., Weaver, A., & Larson, D. (2003). An evaluation of the quantity and quality of empirical research in three pastoral care and counseling journals, 1990–1999: Has anything changed? *The Journal of Pastoral Care & Counseling, 57*(2), 167–178.
- Flannelly, K., Weaver, A., & Handzo, G. (2003). A three-year study of chaplains' professional activities at Memorial Sloan-Kettering Cancer Center in New York City. *Psycho-Oncology, 12*, 760–768.
- Flannelly, K., Weaver, A., & Handzo, G. (2004). A three-year study of chaplains' professional activities at Memorial Sloan-Kettering Cancer Center in New York City. *Chaplaincy Today, 20*(1), 3–12.
- Flannelly, K. J., Galek, K., Bucchino, J., Handzo, G. F., & Tannenbaum, H. P. (2005). Department directors' perceptions of the roles and functions of hospital chaplains: A national survey. *Hospital Topics, 83*(4), 19–27.

- Flannelly, K. J., Galek, K., Bucchino, J., & Vane, A. (2006). The relative prevalence of various spiritual needs. *Scottish Journal of Healthcare Chaplaincy*, 9(2), 25–30.
- Flannelly, K. J., Galek, K., & Flannelly, L. T. (2006). A test of the factor structure of the patient spiritual needs assessment scale. *Holistic Nursing Practice*, 20(4), 187–190.
- Flannelly, K. J., Galek, K., Tannenbaum, H. P., & Handzo, G. F. (2007). A preliminary proposal for a scale to measure the effectiveness of pastoral care with family members of hospitalized patients. *Journal of Pastoral Care and Counseling*, 61(1–2), 19–29.
- Flannelly, K. J., Handzo, G., Weaver, A. J., & Smith, W. J. (2005). A national survey of health care administrators' views on the importance of various chaplain roles. *Journal of Pastoral Care and Counseling*, 59(1–2), 87–96.
- Flannelly, K. J., Handzo, G. F., Galek, K., Weaver, A. J., & Overvold, J. A. (2006). A national survey of hospital directors' views about the importance of various chaplain roles: differences among disciplines and types of hospitals. *Journal of Pastoral Care and Counseling*, 60(3), 213–225.
- Flannelly, K. J., Handzo, G. F., & Weaver, A. J. (2004). Factors affecting healthcare chaplaincy and the provision of pastoral care in the United States. *Journal Pastoral Care Counsel*, 58(1–2), 127–130.
- Flannelly, K. J., Liu, C., Oppenheimer, J. E., Weaver, A. J., & Larson, D. B. (2003). An evaluation of the quantity and quality of empirical research in three pastoral care and counseling journals, 1990–1999: Has anything changed? *Journal of Pastoral Care and Counseling*, 57(2), 167–178.
- Flannelly, K. J., Oettinger, M., Galek, K., Braun-Storck, A., & Kreger, R. (2009). The correlates of chaplains' effectiveness in meeting the spiritual/religious and emotional needs of patients. *Journal of Pastoral Care and Counseling*, 63(1–2), PDF 9, 1-16.
- Flannelly, K. J., Roberts, S. B., & Weaver, A. J. (2005). Correlates of compassion fatigue and burnout in chaplains and other clergy who responded to the September 11th attacks in New York City. *Journal of Pastoral Care and Counseling*, 59(3), 213–224.
- Florell, J. L. (1973). Crisis intervention in orthopedic surgery: Empirical evidence of the effectiveness of a chaplain working with surgery patients. *Bulletin of the American Protestant Hospital Association*, 29–36.
- Fogg, S. L., Weaver, A. J., Flannelly, K. J., & Handzo, G. F. (2004). An analysis of referrals to chaplains in a community hospital in New York over a seven year period. *Journal of Pastoral Care and Counseling*, 58(3), 225–235.
- Galek, K., Flannelly, K. J., Jacobs, M. R., & Barone, J. D. (2008). Spiritual needs: Gender differences among professional spiritual care providers. *Journal of Pastoral Care and Counseling*, 62(1–2), 29–35.
- Galek, K., Flannelly, K. J., Koenig, H. G., & Fogg, S. L. (2007). Referrals to chaplains: The role of religion and spirituality in healthcare settings. *Mental Health, Religion & Culture*, 10, 363–377.
- Galek, K., Flannelly, K. J., Vane, A., & Galek, R. M. (2005). Assessing a patient's spiritual needs: A comprehensive instrument. *Holistic Nursing Practice*, 19(2), 62–69.

- Galek, K., Vanderwerker, L. C., Flannelly, K. J., Handzo, G. F., Kytte, J., Ross, A. M., & Fogg, S. L. (2009). Topography of referrals to chaplains in the Metropolitan Chaplaincy Study. *Journal of Pastoral Care and Counseling*, 63(1–2), 6–13.
- Gartner, J., Larson, D. B., & Vachar-Mayberry, C. D. (1990). A systematic review of the quantity and quality of empirical research published in four pastoral care journals: 1975–1984. *Journal of Pastoral Care*, 44(2), 115–123.
- Grossoehme, D. H. (2001). Self-reported value of spiritual issues among adolescent psychiatric inpatients. *Journal of Pastoral Care*, 55(2), 139–145.
- Grossoehme, D. H. (2004). Patterns of unmet competencies by first-time candidates for APC board certification. *Chaplaincy Today*, 20(2), 13–16.
- Grossoehme, D. H. (2008). Development of a spiritual screening tool for children and adolescents. *Journal of Pastoral Care and Counseling*, 62(1–2), 71–85.
- Grossoehme, D. H., Cotton, S., & Leonard, A. (2007). Spiritual and religious experiences of adolescent psychiatric inpatients versus healthy peers. *Journal of Pastoral Care and Counseling*, 61(3), 197–204.
- Grossoehme, D. H., VanDyke, R., & Seid, M. (2008). Spirituality's role in chronic disease self-management: Sanctification of the body in families dealing with cystic fibrosis. *Journal of Health Care Chaplaincy*, 15(2), 149–158.
- Handzo, G. F., Flannelly, K. J., Kudler, T., Fogg, S. L., Harding, S. R., Hasan, Y. H., . . . Taylor, B. E. (2008). What do chaplains really do? II. Interventions in the New York Chaplaincy Study. *Journal of Health Care Chaplaincy*, 14(1), 39–56.
- Handzo, G. F., Flannelly, K. J., Murphy, K. M., Bauman, J. P., Oettinger, M., Goodell, E., . . . Jacobs, M. R. (2008). What do chaplains really do? I. Visitation in the New York Chaplaincy Study. *Journal of Health Care Chaplaincy*, 14(1), 20–38.
- Hummel, L., Galek, K., Murphy, K. M., Tannenbaum, H. P., & Flannelly, L. T. (2008). Defining spiritual care: An exploratory study. *Journal of Health Care Chaplaincy*, 15(1), 40–51.
- Iler, W. L., Obenshain, D., & Camac, N. (2001). The impact of daily visits from chaplains on patients with chronic obstructive pulmonary disease (COPD): A pilot study. *Chaplaincy Today*, 17(1), 5–11.
- Jankowski, K. R. B., Handzo, G. F., & Flannelly, K. J. (2011). Testing the efficacy of chaplaincy care. *Journal of Health Care Chaplaincy*, 17(3–4), 100–125.
- Jankowski, K. R. B., Vanderwerker, L. C., Murphy, K. M., Montonye, M., & Ross, A. M. (2008). Change in pastoral skills, emotional intelligence, self-reflection, and social desirability across a unit of CPE. *Journal of Health Care Chaplaincy*, 15(2), 132–148.
- Levy, R. D., & Chan, J. (2006). Spiritual coaching in cancer patients: Treating the spirit as well as the disease. *Journal of Pastoral Care and Counseling*, 60(4), 395–404.
- Moran, M., Flannelly, K., Weaver, A., Overvold, J., Hess, W., & Wilson, J. (2005). A study of pastoral care, referral, and consultation practices among clergy in four settings in the New York City area. *Pastoral Psychology*, 53(3), 255–266.
- O'Connor, T. S., Koning, F., Meakes, E., McLarnon-Sinclair, K., Davis, K., & Loy, V. (2001). Quantity and rigor of qualitative research in four pastoral counseling journals. *Journal of Pastoral Care*, 55(3), 271–280.
- Pfannenstiel, B. R. (2003). Sharing information with our peers: What works, what doesn't work. *Chaplaincy Today*, 19(2), 26–30.

- Piderman, K. M., Marek, D. V., Jenkins, S. M., Johnson, M. E., Buryska, J. F., & Mueller, P. S. (2008). Patients' expectations of hospital chaplains. *Mayo Clinic Proceedings*, 83(1), 58–65.
- Popper, K. R. (1989). *Conjectures and refutations: The growth of scientific knowledge*. New York, NY: Routledge.
- Popper, K. R. (1992). *Logic of scientific discovery*. New York, NY: Routledge.
- Roberts, S., Flannelly, K., Weaver, A., & Figley, C. (2003). Compassion fatigue among chaplains, clergy and other respondents after September 11th. *The Journal of Nervous and Mental Disease*, 191(11), 756–758.
- Spidell, S. (2005). A survey of beliefs and practices in professional chaplaincy. *Chaplaincy Today*, 21(1), 23–29.
- Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics*. New York, NY: Harper Collins.
- Tartaglia, A., & Linyear, A. S. (2000). Organ donation: A pastoral care model. *Journal of Pastoral Care*, 54(3), 277–286.
- Taylor, B. E., Flannelly, K. J., Weaver, A. J., & Zucker, D. J. (2006). Compassion fatigue and burnout among Rabbis working as chaplains. *Journal of Pastoral Care and Counseling*, 60(1–2), 35–42.
- Tibbits, D., Ellis, G., Piramelli, C., Luskin, F., & Lukman, R. (2006). Hypertension reduction through forgiveness training. *Journal of Pastoral Care and Counseling*, 60(1–2), 27–34.
- Trothen, T. J. (2001). Canadian supervised pastoral education—affirmations and ethical queries emerging from a two-year study. *Journal of Pastoral Care*, 55(4), 365–377.
- VandeCreek, L. (2000). How has health care reform affected professional chaplaincy programs and how are department directors responding? *Journal of Health Care Chaplaincy*, 10(1), 7–17.
- VandeCreek, L. (2004). How satisfied are patients with the ministry of chaplains? *Journal of Pastoral Care and Counseling*, 58(4), 335–342.
- VandeCreek, L., Siegel, K., Gorey, E., Brown, S., & Toperzer, R. (2001). How many chaplains per 100 inpatients? Benchmarks of health care chaplaincy departments. *Journal of Pastoral Care*, 55(3), 289–301.
- Vanderwerker, L. C., Flannelly, K. J., Galek, K., Harding, S. R., Handzo, G. F., Oettinger, M., et al. (2008). What do chaplains really do? III. Referrals in the New York Chaplaincy Study. *Journal of Health Care Chaplaincy*, 14(1), 57–73.
- Weaver, A., Flannelly, L., Strock, A., Krause, N., & Flannelly, K. (2005). The quantity and quality of research on religion and spirituality in four major gerontology journals between 1985 and 2002. *Research on Aging*, 27(2), 119–135.
- Weaver, A. J., Flannelly, K. J., & Liu, C. (2008). Chaplaincy research: Its value, its quality, and its future. *Journal of Health Care Chaplaincy*, 14(1), 3–19.