Addressing Methodological Challenges in Large Communication Datasets: Collecting and Coding Longitudinal Interactions in Home Hospice Cancer Care

Maija Reblin, PhD, **Margaret F Clayton, PhD, APRN**, **Kevin K John, MA**, and **Lee Ellington, PhD**

College of Nursing, University of Utah, 10 South 2000 East, Salt Lake City UT 84112

Abstract

In this paper, we present strategies for collecting and coding a large longitudinal communication dataset collected across multiple sites, consisting of over 2000 hours of digital audio recordings from approximately 300 families. We describe our methods within the context of implementing a large-scale study of communication during cancer home hospice nurse visits, but this procedure could be adapted to communication datasets across a wide variety of settings. This research is the first study designed to capture home hospice nurse-caregiver communication, a highly understudied location and type of communication event. We present a detailed example protocol encompassing data collection in the home environment, large-scale, multi-site secure data management, the development of theoretically-based communication coding, and strategies for preventing coder drift and ensuring reliability of analyses. Although each of these challenges have the potential to undermine the utility of the data, reliability between coders is often the only issue consistently reported and addressed in the literature. Overall, our approach demonstrates rigor and provides a "how-to" example for managing large, digitally-recorded data sets from collection through analysis. These strategies can inform other large-scale health communication research.

Keywords

methods; health communication; audio-recording; longitudinal; multi-site

Clinical practice, provider education, and patient engagement have each been informed and improved by health communication research findings. For example, the large body of work focusing on the importance of good communication skills for patient outcomes (e.g., Curtis, 2004; Koropchak et al., 2006) has been included in many training programs and curricula, helping new health care providers better communicate with patients during difficult times (Kersun, Gyi, & Morrison, 2009; Rao, Anderson, Inui, & Frankel, 2007; Rosenzweig et al., 2008). Patient communication skills in the form of patient engagement and activation have also been linked to better health outcomes (Hibbard, Mahoney, Stock, & Tusler, 2007; Hibbard, Stockard, Mahoney, & Tusler, 2004; Street, Gordon, Ward, Krupat, & Kravitz,

Corresponding Author: Maija Reblin, PhD, University of Utah, United States.

2005). However, conducting the health communication research that informs these clinical recommendations often presents challenges, especially for large-scale longitudinal studies.

Inherent in any research that relies on recording interactions for communication data collection are concerns that must be addressed to maintain the quality of the data and the resulting research findings. These include making the recording process as unobtrusive as possible for participants, ensuring the maintenance and protection of sensitive data, and ensuring reliability and validity by adequately training coders. While coder reliability is consistently reported in the literature, reporting of strategies to address other factors that may influence the data are inconsistent—leaving future researchers without "best practices" to rely on when developing new recording-based research designs. Thus, we discuss our strategies for addressing methodological challenges, including developing our coding schemes, data collection, data protection and management, and communication coding within the context of our own large-scale, multi-site, longitudinal study of home hospice nurse caregiver interactions.

Context

This article uses as its context a longitudinal observational study conducted as a part of a National Cancer Institute-funded Program Project Grant, Partners in Hospice Care (NCI P01CA138317: PI Mooney), designed to open the "black box" of home hospice nurse-caregiver communication by capturing and describing interactions that have never been systematically studied before (Ellington, Reblin, Clayton, Berry, & Mooney, 2012). Despite greater numbers of family caregivers and increased use of home hospice nationally (National Hospice and Palliative Care Organization, 2013), research on home hospice has not seen similar growth (Tolley & Payne, 2007). In particular, little systematic research has been conducted on effective communication at end-of-life, despite its recognition by patients, caregivers, and health care providers as essential to optimal end-of-life patient care (Cherlin et al., 2005; Fallowfield, Jenkins, & Beveridge, 2002; National Research Council, 2003; Royak-Schaler et al., 2006).

The Nurse-Caregiver Communication study (NCC; Project Leader: Ellington) is one of three R01-type projects within the Partners in Hospice Care Program Project Grant. Across the three studies, a single cohort of over 300 home hospice family caregiver participants, located in 4 diverse locations across the country, is followed from enrollment of the patient into home hospice to patient death and up to 2 years beyond. In addition to NCC, the other projects involve a symptom management intervention and an individually-tailored bereavement support intervention. All three projects share questionnaire and hospice record data; however, NCC is the only one relying on digital audio recordings of visits over time.

Over 225 nurses from participating hospice agencies in our four research locations were asked to audio record their routine in-home communication with hospice caregivers using small digital recorders. These digitally-recorded conversations provide a unique opportunity to study multiple sequential communication interactions in the home with family caregivers. Broader study aims allow us to examine how nurse-caregiver communication processes unfold over time and relate those processes to longitudinally-collected measures of family

caregiver adjustment. This not only contributes data on an understudied population in a naturalistic setting, but the longitudinal nature of our data adds to our ecological validity.

Because of the huge scope of this project—collecting over 2000 hours of audio recorded data from hospice enrollment to patient death from over 300 families and nurses across the country, consisting of over 3000 individual recordings—as well as the challenge of studying communication longitudinally in a difficult-to-recruit population (Demiris, Oliver, Capurro, & Wittenberg-Lyles, 2014; Kutner et al., 2010; McMillan & Weitzner, 2003), it was important for us to develop a protocol that would address potential challenges, including development of coding systems based on our conceptual framework, data collection, data management and protection, and coding, including ensuring and maintaining reliability.

Challenge: Developing Coding Systems Using a Conceptual Framework

Importance & Rationale

An important element to research design is to establish the study and its procedures within a theoretical or conceptual framework. Without thought to the bigger picture, it is difficult to determine the meaning and context of the research findings. To interpret research findings within the context of a larger framework, each measure needs to contribute reliable and valid data to answer a priori research questions within that framework. Often it may be tempting to use measures for the sake of their popularity or because they contain interesting elements, but this potentially contributes to unnecessary participant burden. Our study goals were to capture content and process of communication relevant to domains of care and patient-centered communication. Therefore we sought coding systems that would capture data allowing us to speak to these concepts. For example, though there are many excellent communication systems relevant to couples' conflict discussions, this communication phenomenon is not a primary focus of our research questions, making these coding systems inappropriate for our study. In order to select our coding systems, we first looked at our conceptual framework.

Conceptual Framework

Our conceptual framework was informed both by national clinical practice guidelines and health communication theory. Specifically, our framework was derived from the National Consensus Project for Quality Palliative Care domains of care, and informed by the concept of patient-centered communication (Epstein & Street, 2007; National Consensus Project for Quality Palliative Care, 2013).

Domains of care—The National Consensus Project for Quality Palliative Care has made a major contribution in guiding the advancement of both end-of-life science and care by establishing eight key domains which reflect the delivery of quality palliative care (National Consensus Project for Quality Palliative Care, 2013). We chose to focus on four of the established eight domains of care: physical, psychological, social, and spiritual/existential aspects of care. In addition to being consistent with the hospice philosophy of holistic care and the intersection of our study with the other arms of the program project grant (focused on symptom management and bereavement outcomes), these domains were also consistent

with the home hospice literature, which shows the demands of providing care for an ill and dying significant other often result in caregivers experiencing a reduction in their own physical, psychosocial, and spiritual wellbeing (Chentsova-Dutton et al., 2000; DuBose, 2002; Girgis, Lambert, Johnson, Waller, & Currow, 2013; Grunfeld et al., 2004; Hebert, Arnold, & Schulz, 2007; Kutner et al., 2009; Romito, Goldzweig, Cormio, Hagedoorn, & Andersen, 2013). Thus, we were interested in identifying discussions that incorporated these domains during home hospice visits, and how conversations changed over time as death became imminent.

Patient-centered communication—Because we wanted to capture the important process of dynamic interchange between the caregiver and nurse (Coyle, 2010), we chose Patient Centered Communication to inform our conceptual framework of end of life caregiver nurse communication. Patient centered communication specifies a dynamic interchange between patient and provider and underscores the importance of reciprocal communication. As discussed in an NCI monograph (Epstein & Street, 2007), patient-centered communication consists of the provider's exploration of the patient/caregiver's perspective of health, emotions, and care expectations and a mutual definition of the problem and goals for care. Patient-centered communication has been linked to patient outcomes such as improved emotional states (Coulter, 2012; Zachariae et al., 2003), adherence, improved disease control, physical functioning (Greenfield, Kaplan, & Ware, 1985; Roter, 1977; Rouse & Hamilton, 1991), and reduced health care costs (Stewart et al., 2000). Within a patient-centered approach, the provider and patient establish a working partnership for decision making and goal achievement. Because of the inherent flexibility within patient-centered communication, using this framework allows us to model the ebb and flow of discussions within multiple domains of care over the course end-of-life care.

Communication Coding Development

We used a two-pronged approach to capture both clinical domain content and interpersonal processes of home hospice nurse caregiver communication behaviors. To capture clinical domain content, we developed and validated macro-level approach based on the domains of care in initial pilot work (Ellington, Reblin, Clayton, Berry, & Mooney, 2012; National Consensus Project for Quality Palliative Care, 2013). Our current codebook includes examples derived from this prior work, explanations of each of the domains, and other relevant research literature (e.g. Puchalski, Dorff, & Hendi, 2004). In particular, because of the complexity and nuances of communication associated with the spiritual-religious-existential care domain, we used pilot data to work with experts in palliative care spiritual care to ensure we were accurately recognizing spiritual dialogue (Ellington et al., in press).

Second, to capture interpersonal processes at a more a micro-level, we adapted the Roter Interactional Analysis System (RIAS; Roter & Larson, 2002), one of the most widely used systems designed to code and analyze patient-provider communication. RIAS codes are applied to each utterance, or each individual complete thought, expressed by each speaker. Mutually exclusive and exhaustive categories that reflect the content and process of the dialogue are assigned to these units. Content is distinguished by codes for statements that are primarily informative versus those that are affective. Meanwhile, other codes identify

communication processes, including partnership-building. Furthermore, the system allows coders to indicate who initiates each comment and to whom each comment is directed.

The vast majority of health communication research using RIAS and other similar communication coding systems has been conducted on one-time interactions in traditional clinical interactions between physicians and patients (Kalet, Earp, & Kowlowitz, 1992; Levinson & Roter, 1993; Roter et al., 1995; Roter, Larson, Fischer, Arnold, & Tulsky, 2000). However, some dialogue that occurs in the home simply does not exist in the clinic. For example, because the nurse is entering the family's territory, often partnership building occurs through acknowledging and interacting with objects within the home. There is a potential for increased intimacy with multiple members of the family in home hospice. A nurse may comment on holiday decorations, be offered a snack or recipe, or interact with the family pet (Ellington, Cloyes et al., 2013; Cloyes et al., 2014). Additionally, the home care setting at the end of life often influences care delivery such that it requires more flexibility. Physical assessment can occur inside in a bedroom or outside on a patio. Instruction can be given to caregivers in the kitchen. Moreover, conversations about adjustments of pillows, the placement of furniture, and one's favorite lotion are unlikely to be raised or addressed in the clinic or hospital setting, but may have significance and importance to patients and caregivers in the home environment. Thus, adaptions were made to RIAS to better capture how these conversations develop. These adaptations were tested for feasibility and validity in our pilot work (Ellington, Reblin, Clayton, Berry, & Mooney, 2012).

Traditional application of RIAS has focused on dyadic interactions, whereas often in health care encounters, including and perhaps especially home hospice, multiple individuals are present during nurse visits (Wittenberg-Lyles, Kruse, Oliver, Demiris, & Petroski, 2014). We expanded our scope of coding to include not only the patient, caregiver, and nurse, but also other individuals, such as family or friends. However, a coding challenge emerged from this decision in that at times several distinct conversations took place in the same room simultaneously, often with new voices of unfamiliar participants. Our team dealt with this challenge by prioritizing the conversations in which the nurse participated. Because we collected only audio recordings, there were no visual cues to identify the speaker; however since the nurse wore the recorder, her conversations were often the most clear. Additionally, as a visitor, many times the nurse was introduced to other new individuals in the home. Other strategies to address this coding challenge included singling out and coding single voices with multiple coder passes to listen to the audio recording. This ecologically valid approach is a more accurate representation of health care encounters and could be applied in other settings.

Challenge: Data Collection

Importance & Rationale

Data is the heart of research. How one goes about collecting it can impact the validity of the outcomes; as the old adage says, "Garbage in, garbage out." Ethically, researchers are unable to collect and analyze communication data that is collected in a private location without the consent of participants. This can become challenging when the researchers may not know all the potential participants in an interaction. Research in naturalistic settings, as opposed to

within a lab environment, requires more flexible strategies to ensure the protection and privacy of all individuals. This may be particularly salient in a home health context, in which protected health information is discussed and often intermingled with other sensitive and personal conversations.

A serious risk to validity within the field of health communication is that participants' communication may change in response to obtrusive data collection procedures. Additionally, if data collection becomes overly difficult, individuals may choose not to participate. This is especially true when the researchers are relying on participants to collect their own data. There is often a tradeoff between ecologically valid communication and easing participant burden in collecting communication data. Sending research staff to record each interaction ensures that data is more complete, in that they may be less likely to forget to record and more experienced using recording technology. However, this additional person may change the social dynamic, and therefore the communication within an interaction. Therefore it is important to work out a balance between these elements and develop unobtrusive processes that ensure that the collected communication data remains as natural and complete as possible.

Ethical Considerations

Because our study focused on nurse-caregiver interactions in the home, we consented nurses, caregivers and patients (or their legally-authorized representatives) as participants. However, as mentioned above, others could be present while the nurse was recording, including other family members, friends, or other health care personnel. To address the ethical challenge of consenting what could be a large, changing group of people, we consulted with our Institutional Review Board. To ensure continued consent from the caregivers and patients over multiple visits, as well as to obtain verbal consent from others present, we worked with nurses to make sure they remembered to remind all who were present at the beginning of each visit that she was recording. When new people arrived, nurses were instructed to let them know that she had permission to record from the family and to ask if the new person was comfortable with being recorded. If anyone was uncomfortable, the nurse turned off the recorder. Overwhelmingly, people were willing to be recorded and consistent with other research seemed to continue with normal conversation. All participants were told that they could ask the nurse to discontinue recording at any time or request that a recording be destroyed by the researchers after the fact. Nurses were told to use their clinical judgment about recording during sensitive situations.

Data Collection

A primary challenge for our team was how best to capture sensitive home hospice conversations, balancing the potential for family and provider burden with the capture of clinically valid data, as well as the challenge of portability of data captured in the home. Through collaborative meetings with administrators and nurses from two participating hospices, we developed a detailed protocol for recording procedures (Ellington et al., 2012). In order to limit disruption of clinical care, nurses rather than study staff were responsible for recording home visits. Through our discussions and pilot work, we found that digital recorders worn around the neck of the nurse during visits were unobtrusive and allowed for

quality audio data to be captured while the nurse with a normally conducted home visit. This approach minimized obtrusiveness of data collection and fostered "real world" hospice nurse caregiver interactions.

While audio recording has obvious limitations compared to video recording, we felt that audio recording was more feasible given the technology available. As mentioned, portability had to be considered. Often nurses move throughout the home during their visits; therefore there would be potential limitations to the completeness and quality of video recorded data. Additionally, video recording presents some important privacy concerns beyond audio recording, such as when nurses conduct assessments on private areas of patients' bodies. Further, the use of audio recorded data has been shown to be feasible and comprehensive, with other researchers using audio-recordings to investigate patient-provider communication processes (e.g. Brown, Butow, Boyle, & Tattersall, 2007; Clayton, Dudley, & Musters, 2008; Hausmann et al., 2011; Pollak et al., 2007). Despite the limitations associated with recording audio only, this study is the first of its kind and magnitude, and our systematic large scale examination of audio recorded communication during the provision of home hospice care has the potential to make significant contributions to the end-of-life literature and care, as well as to the development of approaches to ensure the rigor required when managing large longitudinal datasets.

Consenting hospice nurses were given a small Olympus WS-600S digital audio recorder when they were assigned to care for consented caregiver-patient dyads and instructed in their use. Nurses typically kept their recorders for 1 to 2 weeks, recording any visits with study participants during that time. Recorders hold up to approximately 20 hours of audio and were generally collected by study staff during interdisciplinary team meetings, when nurses were already planning to be at the hospice agency. We placed an emphasis on adapting our processes and materials based on nurse feedback—as in the case where improved recorder lanyards were purchased in response to feedback that the older ones were uncomfortable to use. Likewise, the timing and protocol for collection of recorders from hospice nurses varied by agency, and even by individual nurse, depending on agency workflow and nurse preference, ensuring that nurses were not forced to make extra trips to exchange recorders.

Because nurses used the same recorder for all patients over the 1-2 week time period, recording multiple visits over time, and contributing to a pool of over 2000 hours of audio recorded home hospice visits consisting of over 3000 individual recordings, it is important to ensure accurate attribution of a recorded visit to the correct caregiver on the correct date at the time of data collection. As part of initial training in the study protocol, we developed a process where nurses are asked to speak their name, the patient's initials, and visit date into the recorder at the beginning of each visit before they walk up to the patient's door. A reminder card to provide this information was affixed to the recorder pouch, based on nurse feedback. Research staff also kept a log of the dates each nurse had a recorder and programmed the recorders' date and time features, both to be used as a backup to nurse-reported dates. Reported dates were then cross-referenced and verified with the documented visits in the patient medical records.

Data Monitoring

Another challenge when collecting audio recorded data in naturalistic settings is the completion rate of audio-taping by nurses. From a preliminary examination of incoming data, we found some variability in the nurse participant recording rate. Because a nurse's entire patient caseload would not likely participate in our study, recording may be outside of a nurse's typical routine, and therefore difficult to remember. To address this challenge, research staff reminded nurses to record their visits at interdisciplinary team meetings, and through individualized phone calls, emails, and text messages, ranging in frequency from weekly to daily. Mode and frequency of reminder communication was based on nurse preference. The research team evaluated nurse taping rates on a quarterly basis, allowing us to determine which nurses may have needed extra reminders.

Challenge: Data Protection and Management

Importance & Rationale

Data organization strategies need to be considered, tested, and implemented prior to data collection. This is particularly true in large-scale longitudinal studies in which the goal is to analyze communication sequentially. When attempting to track visits over time from many participants across the country, it is unreasonable to assume that the research team would be able to accurately organize data sequentially, or to identify mistakes in the sequence, without an established organizational system to support them. Misidentifying participants, or the order of communication interactions, can have serious implications for the validity of research findings by creating false positive results, as well as obscuring real patterns within the data. Therefore, multiple checks on data management are crucial to maintaining validity. Large-scale studies also need to develop plans for how to prioritize analyses of communication data, due to the time involved in communication coding. Unlike simple quantitative survey data, considerable time is required to analyze each communication interaction, often several times longer than the actual interaction. Because analyzing the entire communication dataset may take a great deal of time, often extending beyond the life of a grant, the research team needs to decide on an analysis approach to ensure that data answering key research questions is analyzed in a timely manner.

With large datasets especially, data security measures must be established and enforced early on. Failure to implement security precautions into data organization systems can prove fatal to a program of research. Breaches of protected health information are associated not only with enormous fines and penalties, but also breaches of trust from the public and damaged reputations for the research team. When participants trust us with insight into their private lives, it is our responsibility to guard it well. This can prove difficult when it requires transporting data across the country and among different members of a research team, but it is essential to master in the face of stiff ethical, legal, and financial repercussions.

Data Protection

Given the sensitivity of end of life conversations, as well as the need to manage multiple lengthy recordings for families in multiple sites across the country, data management and security was a priority, and required a multi-step process. As research staff collected

recorders at various hospice sites, digital audio files were transferred and uploaded via encrypted VPN connection to a secure university server where files are backed up daily and protected by both University and local firewalls. After verified upload, files on the recorders were deleted. Trained research assistants working on secure University computers were responsible for initial data cleaning, including cutting and merging data files and storing the cleaned versions redundantly within the original network folder as well as in a secure file storage online database, also protected by University firewalls and encryption (REDCap Software; Harris et al., 2009). Recordings were deleted if a nurse mistakenly recorded a visit from a non-study participant, or recorded herself testing the recorder. A recording was edited to remove extraneous material if a nurse forgot to stop the recorder after the visit. Extra recording time (unrelated to study hospice visits) ranged from a few minutes to many hours. The enormous size of a multi-hour digital recording is prohibitive for both storage and our analysis software. Further, the recording of unintended interactions may be a privacy concern if the nurse interacts with individuals beyond the scope of the study visit. Thus, our research staff retained only the portion of the recording that contained the study visit. Conversely, the nurse may have stopped the tape at some point during the visit (often this occurred when a nurse identified the visit in her car and stopped the recorder for the trip up to the home for the actual visit), resulting in multiple files for a single visit. In this case, to ensure that all data from a single visit remained together, staff digitally merged the files. All data edits were recorded in study logs.

Data Management

Because of high variability in the number of hospice nurse visits per case, the sheer number of recordings was beyond what prior work and our pilot data indicated. This necessitated the creation of a priority system for initial coding to ensure data most important to address study aims was coded first. As mentioned briefly above, our focus in conducting this research was to observe the communication processes between home hospice nurses and family caregivers in the home, determine how these processes change over the course of hospice enrollment, and make connections between communication and outcomes derived from the other Program Project Grant studies, including symptom management and bereavement outcomes.

Thus, our two initial priorities were to ensure we gathered data from individuals enrolled in multiple studies (not all families participated in all three R01-type projects of the program project grant) and to ensure that we balanced depth within each family's recordings with breadth across different families. We prioritized coding of audio recordings from those families that were participating in all three projects, and developed algorithms to ensure recordings from different points within that family's trajectory (i.e. the first recording, the last recording, the median recording, etc.), which were then randomly assigned to coders. This served to reduce the number of initially coded recordings, particularly among those families that contributed many tapes, while still maintaining good depth for longitudinal analyses, and providing the data for inter-study aims and initial reporting of Program Project results.

Challenge: Communication Coding

Importance & Rationale

The undertaking of communication coding is rife with organizational and training challenges, which must be overcome to establish and maintain reliability and validity. Data collected and managed with the utmost care can devolve into chaos, if coding files are not maintained and associated with any relevant primary data (including original communication data files) in the same careful manner. The human element of communication coding introduces the possibility for errors and differences in judgment, which can potentially impact both the assignment of communication codes and the subsequent research findings. Therefore, adequate training is critical to ensuring valid and reliable coding decisions. Each lab has unique programs for training its coders, whether through classroom experiences, apprenticing with master coders, or working on exemplar data. However, all strategies require coder practice and prompt feedback about validity and reliability of codes, and more complex coding systems require increased training to develop proficient coders. Researchers should account for the time and resource intensiveness of communication coding, and budget for adequate training in their research proposals; otherwise, validity and reliability concerns will undermine the utility of the data.

Coding

Both macro and micro coding systems were accomplished using direct-entry coding software (Noldus Observer XT version 11.5) where codes are assigned based on the audio-recording, as opposed to transcripts, which are not usually financially feasible for large scale studies. Moreover, direct entry coding allows coders to interpret voices and inflections, affording context that may be missing from a transcription alone (e.g. intonation and affect; Roter & Larson, 2002). To facilitate this audio-based coding, basic information about the participants, such as caregiver and patient first names, relationships, and genders, was made available on our secure database for trained coders. Though often other family and friends were involved in the visit, especially across the course of multiple visits, it was important for our study aims to identify the primary hospice-designated caregiver. Based on provided information, coders were better able to identify voices (e.g. an older female voice is a wife, a younger female voice is a daughter).

Once coders completed coding any given recording, files were uploaded to the data management website, from which they became accessible to the project coordinator, who accessed them for storage to the local University server for analysis, including reliability. Coding for both macro and micro approaches occurred independently, but both files for a particular recording could be merged at the time of analysis based on digital time stamps to provide a layered view of communication and allow the comparison of data derived from both approaches. We used separate and independent coding teams for each approach (one team coding for domains of care and another using adapted RIAS for finer grained analysis) to ensure that delays from one group of coders did not hinder the progress of the other team. This system allows each team to specialize and focus on only their specific set of codes, reducing coder burden and resulting in faster coding.

A percentage of our recordings were used for training purposes and to assess reliability. These were double-coded, or coded by multiple individuals. Reliability statistics were calculated on double-coded files for both types of coding systems using Noldus Observer XT. Settings within the program allowed for a small window of discrepancy for the match between timestamps to account for variation in when each coder initiates a particular code (e.g. in the middle or at the end of an utterance). Coding files that met or exceeded the reliability threshold (overall .60 Kappa for micro-level coding using adapted RIAS; 85% agreement for macro-level coding reflecting domains of care) were considered to have good reliability (Bakeman, 2000), and one of the two files for that recording was randomly selected to be included in a master file containing all final coding files. Coders that submitted files that failed to meet the standard were asked to meet with their coding partner to discuss and address coding discrepancies based on feedback from the project coordinator. The project coordinator consistently reviewed files for common coding issues (e.g. confusion between two codes, different number of utterances identified, mismatch of speaker). After discussion, which could be arbitrated by the project coordinator or study investigators in difficult situations, coders prepared a consensus file, which was then checked for resolution of conflict, and included in the master file.

Weekly coder meetings were held for each group to assess validity and reliability, as well as to prevent coder drift. Beginning during coder training and at coder meetings, we discussed challenges with specific recordings, and invited coders to bring up particular scenarios or issues that they found interesting or difficult to code. This led to frequent discussions of the definitions and limitations of our coding system. Maintaining the validity of the coding scale is a continual process throughout the duration of the project. We were also mindful throughout the study of the emotional aspects inherent in our data, including witnessing the decline of a patient, expressions of pain or extreme grief, or conversations that strike a personal note with a given coder. We regularly discussed these emotional aspects of coding as a way of supporting our coders who listen to these difficult conversations on a regular basis as part of their coding responsibilities (Ellington, Reblin, Berry, Giese-Davis, & Clayton, 2013; Shakespeare et al., 2007).

Discussion

Obtaining and managing large digitally recorded data sets is challenging, requiring a highly detailed plan for developing a coding system, data collection, data management and protection, and data analysis. However, these challenges can be overcome, facilitating contextually relevant "real world" health communication research and application to clinical practice. In this paper, we shared our solutions to challenges within the context of our longitudinal multi-site multi-state study investigating how over 300 designated family caregivers worked with their hospice nurses to care for patients during the last weeks of life.

In our study, we use a clinical and conceptual -based approach to represent communication behaviors. Based on existing frameworks of Domains of Care and Patient-Centeredness, we developed and adapted macro and micro levels of coding to identify both content and process of communication. We also developed tailored protocols for data collection to ensure good recording compliance. Data management and protection included protocols for

secure data cleaning and storage. Communication coding was completed by independent coding teams using direct-entry coding software which improved contextual relevance and validity based on affective tone, in addition to conserving research funds. Regular coder meetings and assessment of coded data ensured that we maintained reliability and avoided coder drift. Overall, our approach demonstrates rigor, providing an example of how to manage large audio-recorded data sets from data collection through data analysis. This procedure for capturing ecologically valid health communication data could be adapted to many types of health communication research.

Because of the large scale, longitudinal nature, and ecological validity of studies such as ours, very rich communication data can be obtained. However, this comes at the cost of extensive effort and resources, which may not be available to all researchers. To combat this, and to more fully honor the contribution of our participants, researchers and scholars in partnership with IRBs need to find expedient ways to share data from such large scale studies once collected. In this way multiple research questions from different perspectives can be answered and disseminated into the field. For example, given the large scale nature of our current study, our primary aims focus only on quantitative coding schemes. However, we have developed several ongoing partnerships with qualitative researchers from different disciplines to widen the reach of our dataset (Cloyes, Berry, Reblin, Clayton, & Ellington, 2012; Tjia, Reblin, LeMay, Clayton, & Ellington, 2014). The richness of our data lends itself to parceling out smaller scale qualitative projects addressing unique research questions. Not only do multiple fields of research benefit from these interdisciplinary collaborations, but working with these partners informs our own perspectives and future work.

Although the scope of this paper was on data collection and management, a further methodological complexity should be mentioned as a frequent limitation in communication research: the focus on a single dyad (Kreps & Viswanath, 2001; O'Hair, 2003; Roter & Hall, 1997), to the exclusion of the larger healthcare team. Almost all health care, and particularly home hospice care, is provided by a team that includes multi-disciplinary health care professionals and family caregivers, yet we have little understanding about and few evidence-based recommendations for how these teams communicate with each other and with the patient and their family (Dawson, 2007; Street & Blackford, 2001; Wallace, 2001; Wittenberg-Lyles et al., 2013). This is also a limitation of our own study; as a first step, we chose to record only the hospice nurse case manager during interaction with the caregiver. However, future work should include team interactions to more accurately model the true health care experience, with special focus on how these interactions change over time.

For those studies that do truly focus on dyadic interactions, attention should be given to appropriate statistical analysis. Few studies in the literature actually conduct true dyadic analysis that accounts for the non-independence of the communication participants (Griffin & Gonzalez, 2003; Li & Loke, 2014; Manne & Badr, 2008). The progress of statistical packages has facilitated multilevel modeling and structural equation modeling, but researchers have been slow to take advantage of these tools to present more accurate and robust analyses of their data. Using appropriate statistical tools when analyzing data collected from large-scale ecologically valid research, we have the opportunity to determine

what communication skills and patterns are associated with better patient, family, and health care provider outcomes, and by extension influence practice.

Health communication is complex, and undertaking the study of these interactions in meaningful and systematic ways presents many methodological challenges. We hope that by addressing some of the complexities that accompany large-scale more ecologically valid health communication studies, we have provided solutions others might employ. Across health care fields, health communication researchers have the potential to contribute contextually relevant and important research to a growing evidence base for clinical intervention and application.

Acknowledgments

References

Bakeman, R. Behavioral observation and coding. In: Reis, HT.; Judd, CM., editors. Handbook of research methods in social and personality psychology. New York: Cambridge University Press; 2000. p. 138-159.

Brown RF, Butow PN, Boyle F, Tattersall MHN. Seeking informed consent to cancer clinical trials: Evaluating the efficacy of doctor communication skills training. Psycho-Oncology. 2007; 16:507–516. [PubMed: 16986176]

Chentsova-Dutton Y, Shuchter S, Hutchin S, Strause L, Burns K, Zisook S. The psychological and physical health of hospice caregivers. Annals of Clinical Psychiatry. 2000; 12:19–27. [PubMed: 10798822]

Cherlin E, Fried T, Prigerson HG, Schulman-Green D, Johnson-Hurzeler R, Bradley EH. Communication between physicians and family caregivers about care at the end of life: When do discussions occur and what is said? Journal of Palliative Medicine. 2005; 8:1176–1185. [PubMed: 16351531]

Clayton MF, Dudley WN, Musters A. Communication with breast cancer survivors. Health Communication. 2008; 23:207–221. [PubMed: 18569050]

Cloyes KG, Berry PH, Reblin M, Clayton M, Ellington L. Exploring communication patterns among hospice nurses and family caregivers: A content analysis of in-home speech interactions. Journal of Hospice & Palliative Nursing. 2012; 14:426–437.

Cloyes KG, Carpenter JG, Berry PH, Reblin M, Clayton MF, Ellington L. "A true human interaction": Comparison of family caregiver and hospice nurse perspectives on needs of family hospice caregivers. Journal of Hospice & Palliative Nursing. 2014; 16:282–290.

Coulter A. Patient engagement—What works? The Journal of Ambulatory Care Management. 2012; 35:80–89. [PubMed: 22415281]

Coyle, N. Introduction to palliative care nursing care. In: Ferrell, BR.; Coyle, N., editors. Oxford textbook of palliative nursing. 3rd. New York: Oxford University Press; 2010. p. 3-11.

Curtis JR. Communicating about end-of-life care with patients and families in the intensive care unit. Critical Care Clinics. 2004; 20:363–380. [PubMed: 15183208]

Dawson S. Interprofessional working: Communication, collaboration… perspiration! International Journal of Palliative Nursing. 2007; 13:502–505. [PubMed: 18073710]

Demiris G, Oliver DP, Capurro D, Wittenberg-Lyles E. Implementation science: Implications for intervention research in hospice and palliative care. The Gerontologist. 2014; 54:163–171. [PubMed: 23558847]

DuBose ER. A special report: Spiritual care at the end of life. Second Opinion. 2002; 10:4–74. [PubMed: 16025645]

Ellington L, Cloyes K, Berry P, Thomas NT, Reblin M, Clayton MF. Complexities for hospice nurses in supporting family caregivers: Opinions from US thought leaders. Journal of Palliative Medicine. 2013; 16:1013–1019. [PubMed: 23837448]

Ellington L, Reblin M, Berry P, Giese-Davis J, Clayton MF. Reflective research: Supporting researchers engaged in analyzing end-of-life communication. Patient Education & Counseling. 2013; 91:126–128. [PubMed: 23092615]

Ellington L, Reblin M, Clayton MF, Berry P, Mooney K. Hospice nurse communication with cancer patients and their family caregivers. Journal of Palliative Medicine. 2012; 15:262–268. [PubMed: 22339285]

Ellington L, Reblin M, Ferrell B, Puchalski C, Otis-Green S, Handzo G, Doyan K, Clayton MF. The religion of 'I don't know': Naturalistic pilot observations of spiritual conversations occurring during cancer home hospice nurse visits. Omega: Journal of Death & Dying. in press.

Epstein, RM.; Street, RL, Jr. Patient-centered communication in cancer care: Promoting healing and reducing suffering. Bethesda, MD: National Cancer Institute; 2007. NIH Report No. 07-6225

Fallowfield LJ, Jenkins VA, Beveridge HA. Truth may hurt but deceit hurts more: Communication in palliative care. Palliative Medicine. 2002; 16:297–303. [PubMed: 12132542]

Girgis A, Lambert S, Johnson C, Waller A, Currow D. Physical, psychosocial, relationship, and economic burden of caring for people with cancer: A review. Journal of Oncology Practice. 2013; 9:197–202. [PubMed: 23942921]

Greenfield S, Kaplan S, Ware JE Jr. Expanding patient involvement in care. Effects on patient outcomes. Annals of Internal Medicine. 1985; 102:520–528. [PubMed: 3977198]

Griffin D, Gonzalez R. Models of dyadic social interaction. Philosophical Transactions of the Royal Society of London Series B: Biological Sciences. 2003; 358:573–581. [PubMed: 12689382]

Grunfeld E, Coyle D, Whelan T, Clinch J, Reyno L, Earle CC, Willan A, Viola R, Coristine TJ, Glossop R. Family caregiver burden: Results of a longitudinal study of breast cancer patients and their principal caregivers. Canadian Medical Association Journal. 2004; 170:1795–1801. [PubMed: 15184333]

Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. Journal of Biomedical Informatics. 2009; 42:377–381. [PubMed: 18929686]

Hausmann LR, Hannon MJ, Kresevic DM, Hanusa BH, Kwoh CK, Ibrahim SA. Impact of perceived discrimination in healthcare on patient-provider communication. Medical Care. 2011; 49:626–633. [PubMed: 21478769]

Hebert RS, Arnold RM, Schulz R. Improving well-being in caregivers of terminally ill patients: Making the case for patient suffering as a focus for intervention research. Journal of Pain and Symptom Management. 2007; 34:539–546. [PubMed: 17616333]

Hibbard JH, Mahoney ER, Stock R, Tusler M. Do increases in patient activation result in improved self-management behaviors? Health Services Research. 2007; 42:1443–1463. [PubMed: 17610432]

Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM): Conceptualizing and measuring activation in patients and consumers. Health Services Research. 2004; 39:1005–1026. [PubMed: 15230939]

Kalet A, Earp JA, Kowlowitz V. How well do faculty evaluate the interviewing skills of medical students? Journal of General Internal Medicine. 1992; 7:499–505. [PubMed: 1403205]

Kersun L, Gyi L, Morrison WE. Training in difficult conversations: A national survey of pediatric hematology–oncology and pediatric critical care physicians. Journal of Palliative Medicine. 2009; 12:525–530. [PubMed: 19508138]

Koropchak CM, Pollak KI, Arnold RM, Alexander SC, Skinner CS, Olsen MK, Jeffreys AS, Rodriguez KL, Abernethy AP, Tulsky JA. Studying communication in oncologist-patient encounters: The SCOPE trial. Palliative Medicine. 2006; 20:813–819. [PubMed: 17148536]

Kreps G, Viswanath K. Communication interventions and cancer control: A review of the National Cancer Institute's health communication intervention research initiative. Family & Community Health. 2001; 24:ix–xiii. [PubMed: 11589219]

Kutner J, Kilbourn KM, Costenaro A, Lee CA, Nowels C, Vancura JL, Anderson D, Keech TE. Support needs of informal hospice caregivers: A qualitative study. Journal of Palliative Medicine. 2009; 12:1101–1104. [PubMed: 19764830]

Kutner J, Smith M, Mellis K, Felton S, Yamashita T, Corbin L. Methodological challenges in conducting a multi-site randomized clinical trial of massage therapy in hospice. Journal of Palliative Medicine. 2010; 13:739–744. [PubMed: 20597707]

Levinson W, Roter D. The effects of two continuing medical education programs on communication skills of practicing primary care physicians. Journal of General Internal Medicine. 1993; 8:318–324. [PubMed: 8320576]

Li Q, Loke AY. A literature review on the mutual impact of the spousal caregiver-cancer patients dyads: 'Communication', 'reciprocal influence', and 'caregiver-patient congruence'. European Journal of Oncology Nursing. 2014; 18:58–65. [PubMed: 24100089]

Manne S, Badr H. Intimacy and relationship processes in couples' psychosocial adaptation to cancer. Cancer. 2008; 112:2541–2555. [PubMed: 18428202]

McMillan SC, Weitzner MA. Methodologic issues in collecting data from debilitated patients with cancer near the end of life. Oncology Nursing Forum. 2003; 30:123–129. [PubMed: 12515990]

O'Hair D. Research traditions in provider-consumer interaction: Implications for cancer care. Patient Education & Counseling. 2003; 50:5–8. [PubMed: 12767576]

National Consensus Project for Quality Palliative Care. Clinical practice guidelines for quality palliative care. 3rd2013. http://www.nationalconsensusproject.org

National Hospice and Palliative Care Organization. NHPCO's facts and figures: Hospice care in America. Alexandria, VA: National Hospice and Palliative Care Organization; 2013.

National Research Council. Describing death in America: What we need to know. Washington, D.C: National Academies Press; 2003.

Pollak KI, Arnold RM, Jeffreys AS, Alexander SC, Olsen MK, Abernethy AP, Skinner CS, Rodriguez KL, Tulsky JA. Oncologist communication about emotion during visits with patients with advanced cancer. Journal of Clinical Oncology. 2007; 25:5748–5752. [PubMed: 18089870]

Puchalski CM, Dorff RE, Hendi IY. Spirituality, religion, and healing in palliative care. Clinics in Geriatric Medicine. 2004; 20:689–714. [PubMed: 15541620]

Rao JK, Anderson LA, Inui TS, Frankel RM. Communication interventions make a difference in conversations between physicians and patients: A systematic review of the evidence. Medical Care. 2007; 45:340–349. [PubMed: 17496718]

Romito F, Goldzweig G, Cormio C, Hagedoorn M, Andersen BL. Informal caregiving for cancer patients. Cancer. 2013; 119:2160–2169. [PubMed: 23695928]

Rosenzweig M, Hravnak M, Magdic K, Beach M, Clifton M, Arnold R. Patient communication simulation laboratory for students in an acute care nurse practitioner program. American Journal of Critical Care. 2008; 17:364–372. [PubMed: 18593836]

Roter DL. Patient participation in the patient-provider interaction: The effects of patient question asking on the quality of interaction, satisfaction and compliance. Health Education & Behavior. 1977; 5:281–315.

Roter, DL.; Hall, JA. Patient-provider communication. In: Glanz, K.; Lewis, FM.; Rimer, BK., editors. Health behavior and health education: Theory, education, and practice. 2nd. San Francisco: Jossey-Bass; 1997. p. 206-226.

Roter DL, Hall JA, Kern DE, Barker LR, Cole KA, Roca RP. Improving physicians' interviewing skills and reducing patients' emotional distress: A randomized clinical trial. Archives of Internal Medicine. 1995; 155:1877–1884. [PubMed: 7677554]

Roter DL, Larson S, Fischer GS, Arnold RM, Tulsky JA. Experts practice what they preach: A descriptive study of best and normative practices in end-of-life discussions. Archives of Internal Medicine. 2000; 160:3477–3485. [PubMed: 11112242]

Roter DL, Larson S. The Roter Interaction Analysis System (RIAS): Utility and flexibility for analysis of medical interactions. Patient Education & Counseling. 2002; 46:243–251. [PubMed: 11932123]

Rouse RA, Hamilton MA. Dentists evaluate their patients: An empirical investigation of preferences. Journal of Behavioral Medicine. 1991; 14:637–648. [PubMed: 1791626]

Royak-Schaler R, Gadalla SM, Lemkau JP, Ross DD, Alexander C, Scott D. Family perspectives on communication with healthcare providers during end-of-life cancer care. Oncology Nursing Forum. 2006; 33:753–760. [PubMed: 16858457]

Shakespeare TP, Tang JI, Shen L, Lu JJ, Mukherjee RK, Lee KM, Wynne CJ, Back MF. Does the implementation of radiation oncology outpatient infection control measures adversely affect patient satisfaction with doctor-patient interaction? Singapore Medical Journal. 2007; 48:246–251. [PubMed: 17342296]

Stewart M, Brown JB, Donner A, McWhinney IR, Oates J, Weston WW, Jordan J. The impact of patient-centered care on outcomes. The Journal of Family Practice. 2000; 49:796–804. [PubMed: 11032203]

Street A, Blackford J. Communication issues for the interdisciplinary community palliative care team. Journal of Clinical Nursing. 2001; 10:643–650. [PubMed: 11822515]

Street RL Jr, Gordon HS, Ward MM, Krupat E, Kravitz RL. Patient participation in medical consultations: Why some patients are more involved than others. Medical Care. 2005; 43:960–969. [PubMed: 16166865]

Tjia, J.; Reblin, M.; LeMay, C.; Clayton, MF.; Ellington, L. Organization, teamwork and advocacy: Important skills needed by hospice family caregivers who manage medications; Paper presented at the ASCO Palliative Care in Oncology Symposium; Boston, MA. 2014 Oct.

Tolley DC, Payne R. Hospice partnerships with academic entities: Philosophical and historical background and assessment of future needs. Journal of Pain & Symptom Management. 2007; 33:90–98. [PubMed: 17196910]

Wallace PR. Improving palliative care through effective communication. International Journal of Palliative Nursing. 2001; 7:86–90. [PubMed: 12271254]

Wittenberg-Lyles E, Kruse RL, Oliver DP, Demiris G, Petroski G. Exploring the collective hospice caregiving experience. Journal of Palliative Medicine. 2014; 17:50–55. [PubMed: 24351126]

Wittenberg-Lyles E, Oliver DP, Kruse RL, Demiris G, Gage LA, Wagner K. Family caregiver participation in hospice interdisciplinary team meetings: How does it affect the nature and content of communication? Health Communication. 2013; 28:110–118. [PubMed: 22435889]

Zachariae R, Pedersen CG, Jensen AB, Ehrnrooth E, Rossen PB, Von Der Maase H. Association of perceived physician communication style with patient satisfaction, distress, cancer-related self-efficacy, and perceived control over the disease. British Journal of Cancer. 2003; 88:658–665. [PubMed: 12618870]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript