Strategies to Maximize Data Collection Response Rates in a Randomized Control Trial Focused on Children with Medical Complexity

Stanley M. Finkelstein, PhD,[1] Margaret Celebrezze, MA, MPH,[1,2] Rhonda Cady, PhD,[1,3] Scott Lunos, MS,[4] and Wendy S. Looman, PhD[2]

[1]Department of Laboratory Medicine and Pathology/Health Informatics, [2]School of Nursing, and [4]Clinical and Translational Science Institute, University of Minnesota, Minneapolis, Minnesota.
[3]Gillette Children's Specialty Healthcare, St. Paul, Minnesota.

Abstract

Introduction: Obtaining complete and timely subject data is key to the success of clinical trials, particularly for studies requiring data collected from subjects at home or other remote sites. A multifaceted strategy for data collection in a randomized controlled trial (RCT) focused on care coordination for children with medical complexity is described. The influences of data collection mode, incentives, and study group membership on subject response patterns are analyzed. Data collection included monthly healthcare service utilization (HCSU) calendars and annual surveys focused on care coordination outcomes. Materials and Methods: One hundred sixty-three families were enrolled in the 30-month Tele-Families RCT. Subjects were 2–15 years of age at enrollment. HCSU data were collected by parent/guardian self-report using mail, e-mail, telephone, or texting. Surveys were collected by mail. Incentives were provided for completed surveys after 8 months to improve collection returns. Outcome measures were the number of HCSU calendars and surveys returned, the return interval, data collection mode, and incentive impact. Results: Return rates of 90% for HCSU calendars and 82% for annual surveys were achieved. Mean return intervals were 72 and 65 days for HCSU and surveys, respectively. Survey response increased from 55% to 95% after introduction of a gift card and added research staff. Conclusions: High return rates for HCSU calendars and health-related surveys are attainable but required a flexible and personnel-intensive approach to collection methods. Family preference for data collection approach should be obtained at enrollment, should be modified as needed, and requires flexible options, training, intensive staff/family interaction, and patience.

Key words: *home health monitoring, telehealth, telenursing, information management*

Introduction

Akey element in the success of a clinical trial is the completeness and timeliness by which subject data are obtained by the trial investigators.[1–4] Although this is generally not an issue for data collected "in person" during clinic visits, data collected via mail, e-mail, or telephone are more likely to be delayed or incomplete, which negatively impacts external validity and generalizability. The aims of this article are to describe multifaceted strategies for collecting data in a clinical trial and to analyze the influence of these strategies and study group membership on subject data response rates and return intervals.

A large volume of literature details the problems and potential solutions associated with the data collection effort.[5–9] Dillman[5] addressed measurement and nonresponse errors in his text on the Tailored Design Method. Dillman's approach, based on Social Exchange Theory, provides specific guidance on survey design and implementation that increased the timeliness and response rate to self-administered surveys from study participants.

Social Exchange Theory emphasizes that people are likely to engage in behavior they find rewarding and avoid behavior they find costly.[10] In this context cost reduction may involve easier survey completion because of improvements in study instruments (less complex, fewer items, less ambiguous language, improved physical presentation, etc.) and flexible submission methods that do not interfere with the family's daily routines.

Additional factors that may impact survey response rates were detailed in several healthcare studies using postal questionnaires,[7] alternative modalities,[6] healthcare professional subjects,[8] and families under stress.[9] These examples focused on randomized controlled trials (RCTs) to evaluate different methods to improve response rates in survey data

collection. Results from these studies were mixed regarding the benefits of different strategies to increase reporting rates.[7] This article is the first to report on the use of a broad range of evidence-based methods to maximize data collection in a clinical RCT focused on care coordination for children with medical complexity.

Materials and Methods

SETTING AND SUBJECTS

The use of multifaceted methods for enhancing data collection was a key element of the TeleFamilies research program.[11] TeleFamilies was a three-armed RCT testing the effectiveness of advanced practice registered nurse (APRN) care coordination for children with medical complexity.[12] The study was based in a large, urban, general pediatrics primary-care clinic affiliated with a nonprofit children's hospital. The clinic used a medical home model of care coordination with the primary care provider managing overall care.[13] Children recruited for this study were between 2 and 15 years of age, from English-speaking households, and satisfied four of the five conditions specified in the Children with Special Healthcare Needs Screener®.[14] Those agreeing to participate provided written informed consent, following the guidelines of the Institutional Review Boards.

The 163 consented subjects were randomized into one of three study arms. A stratified block randomization was used, with stratification based on age. A control group received registered nurse (RN)/licensed practical nurse (LPN)–delivered telephone triage and team-based care coordination. Two intervention groups were assigned a single clinic-based APRN care coordinator who provided a consistent point of contact and who worked collaboratively with each child's primary care provider before, during, after, and between clinic visits.[15]

APRN-delivered care coordination was conducted primarily via telehealth. Parents of subjects in the telephone intervention group used telephone telehealth to conduct these activities, whereas parents of subjects in the video intervention group used telephone and interactive clinic–home video. Video visits used a Health Insurance Portability and Accountability Act–compliant, Web-based platform (Virtual Interactive Families©, Cedar Falls, IA).

Contact with families in the control group was initiated by parents or caregivers, with follow-up calls initiated by the triage nurses as needed to complete a delegated task or provide requested information or advice.

A primary difference between the control and intervention groups was the intensity and focus of the interactions with the nurse. In the control group, encounters were protocol-based and often limited to the exchange of information. In the intervention groups, encounters were often proactive and included assessment and advanced clinical management by the APRN. The APRN's maximum caseload was 105 families, whereas the triage nurses responded to calls for all clinic patients in addition to those in the study. Nurse interactive contacts were based on need and the different concerns that the RN/LPN for the controls and the APRN for the intervention groups could address, so that subjects in each study group did not receive the same number of contacts. Similar to RN/LPN services, APRN care coordination was available during weekday work hours. All after-hours requests rolled over to a standard after-hours triage service. Study participation encompassed an initial 6-month period to document baseline healthcare service utilization patterns, followed by a 2-year intervention period, for a total 30 months of data collection for each subject.

DATA COLLECTION

Outcome measurements included healthcare service utilization (HCSU) data and an annual survey. HCSU data were collected monthly by self-report from families using a project-specific HCSU calendar form (*Fig. 1*). The annual survey, collected from study families by mail, consisted of self-reporting the items in the PedsQL 4.0–Core,[16] PedsQL 3.0–Parent Family Impact,[17] Consumer Assessment of Health Plan Satisfaction,[18] DISC (help discrepancy),[19] Functional Status II®,[20] and TMPQ (telehealth perception) measures.[21] There were 146 questions in the survey packet; time needed to complete a survey packet was approximately 60 min. Survey data were collected at the conclusion of the baseline period (i.e., enrollment Month 6), and annually during the intervention period (i.e., at enrollment Months 18 and 30).

HCSU DATA COLLECTION

An instructional letter with 3 months of HCSU calendars to use for tracking their child's healthcare utilization (i.e., appointments and visits) was sent to subject families. They were asked to return the completed calendars at the end of each month using pre-addressed, stamped envelopes. Families were called 30 days after enrollment to confirm receipt of the forms and address questions or concerns. Receipt of completed HCSU forms was tracked monthly. Families not returning forms received a reminder telephone call. If no one answered, a voicemail message was left, and a follow-up call was scheduled.

Family comments during telephone reminders uncovered information that influenced future HCSU data collection strategies. Families use a variety of mechanisms to track their child's healthcare appointments, such as a cell phone or work

Fig. 1. Healthcare service utilization calendar.

calendar, and often preferred reporting these data by telephone, e-mail, or text message rather than completing and mailing the paper form. Based on this feedback and ongoing response rate review, the HCSU data collection process was expanded to include these strategies. Family HCSU data collection preferences were noted along with contact timing preferences (e.g., "Mom works the night shift; call after 3:00 p.m.").

Despite expanded strategies, some families often forgot to send the monthly HCSU data, necessitating additional reminders from study personnel. If 3 or more months of unsuccessful reminders elapsed, appointment information was abstracted from the subject's electronic medical record (EMR). The EMR-obtained data were confirmed with the family whenever possible. EMR confirmation assisted recall of prior

month appointments, determined whether work or school was missed, and captured visit information not documented in the EMR, such as out-of-network specialty appointments and therapy visits. Unconfirmed HCSU data were noted as collected from EMR.

SURVEY DATA COLLECTION

Initial survey data collection efforts incorporated the approach of Dillman.[5] A baseline survey and a pre-addressed, stamped return envelope were mailed to all subject families during their fifth month of enrollment. Survey mailings at Year 1 and Year 2 were personalized and made more visually attractive. Following Dillman's recommendations this included utilization of colorful cover images on the packet envelopes, handwritten addresses, personalized

cover letters, and the use of stamps instead of metered postage. A colorful reminder and/or a thank-you postcard was sent approximately 2 weeks after the surveys were mailed. Reminders continued to be sent to nonresponders on a monthly basis for up to 3 months. Timing of survey mailing considered potential barriers that may distract the family from completing the survey, such as start of the school year, holidays, and Mondays, when larger amounts of mail might be received.

Similar to HCSU data collection processes, changes in survey collection processes were initiated based on caregiver feedback and insights gained from working directly with study families. Modifications included an anticipatory "heads up" telephone call prior to mailing of survey packets, contact with the family if surveys were not returned within 30 days, face-to-face reminders during scheduled clinic appointments, and providing replacement surveys for families to complete during clinic appointments or at home. Upon request, study personnel assisted families with completing the survey in clinic or in their home. This proved helpful for caregivers with vision problems and with English language written literacy challenges, as well as those caring for additional children who accompanied them to clinic appointments. Additional modifications included a bolded "return by…" statement in the cover letter and reformatting the survey into a more attractive and easier-to-read booklet. Two monetary incentives were added in an effort to improve survey return rates. A $2 bill was included with each survey mailing starting in study Month 14, and after a 6-month test period the incentive shifted to a $20 gift card from a major retailer upon return of a completed survey booklet.

DATA ANALYSIS

Descriptive and statistical analysis used SAS version 9.3 software (SAS Institute Inc., Cary, NC). The primary outcome was the number of completed data instruments returned (monthly HCSU calendars and annual surveys). Additional outcomes were the return interval (number of days between the request for and receipt of completed HCSU calendars and surveys), data collection mode (i.e., mail, e-mail, telephone, in-person, or text) for HCSU, and incentive type in place upon project receipt of the survey.

Logistic regression models with a random subject effect compared the proportion of each data collection mode between study groups. Fisher's exact test compared the primary mode of collection for each person between groups. The primary collection mode is the mode most used by each subject for HCSU response/

confirmation. A general linear model with a random subject effect compared the return interval between groups and between data collection modes. The Kruskal–Wallis test was used for differences between return intervals for the surveys. Values of $p < 0.05$ were considered statistically significant.

Results

There were 163 subject families enrolled at the start of the TeleFamilies trial. Over the course of the study, 14 subject families withdrew, and 1 subject died, resulting in 148 subject families completing all 30 months of study participation.

HCSU RESULTS

Of the total of 4,697 HCSU calendars, 4,164 were collected from subjects over the 30-month study period: 1,294 (84%) from the control group, 1,379 (88%) from the telephone group, and 1,491 (94%) from the video group. *Figure 2* shows the influence of the five data collection modes on subject responses in each arm of the study during the 2-year intervention period. The 6-month baseline period data collection was not included because alternative data collection modes were not yet implemented. There were no significant differences among study groups for each data collection mode and no such differences in primary collection mode choice across the study groups. However, between-group comparison of response rates was significantly different ($p = 0.014$).
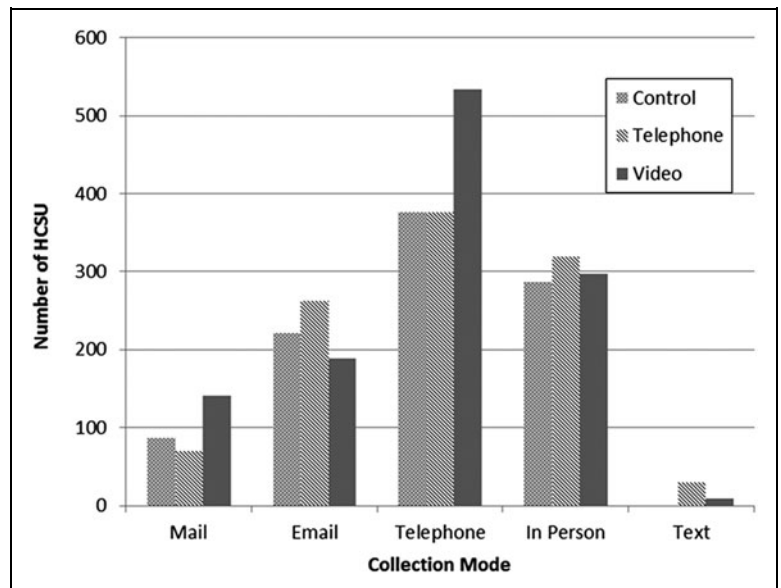

Fig. 2. Total number of healthcare service utilization (HCSU) forms returned, by study arm and data collection method, during the 2-year intervention period ($n = 3,201$). Baseline data are not reported here.
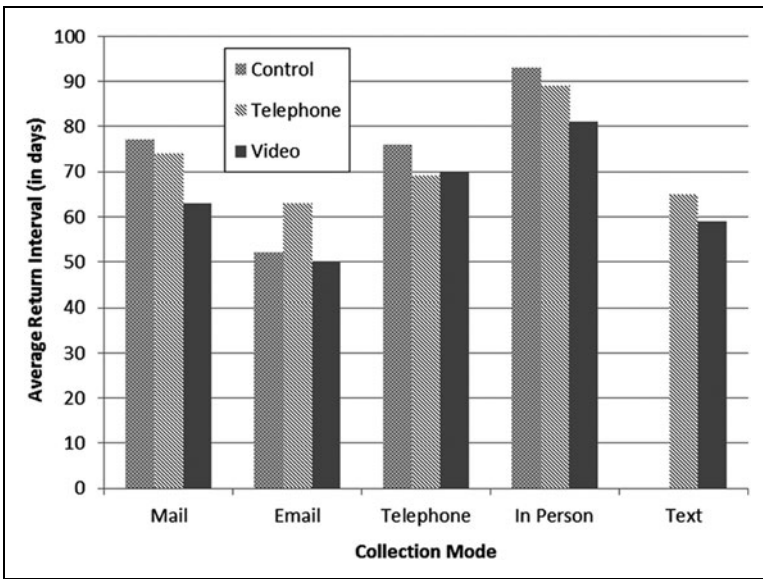
Fig. 3. Healthcare service utilization return intervals for all study groups, by data collection method, during intervention Years 1 and 2.

HCSU return intervals are shown in *Figure 3*. Mean (standard deviation) return intervals were 76 (55), 74 (60), and 69 (46) days for the control, telephone, and video groups, respectively. The overall mean return interval for all subjects via all collection methods was 72 (54) days. A significant difference in mean return intervals between the collection modes ($p < 0.0001$) was found. In pairwise comparisons e-mail had a shorter return interval than telephone, mail, and in-person ($p < 0.0001$), and telephone had a shorter return interval than in-person collection ($p < 0.0006$). The interaction between study group and collection mode for return intervals was not statistically significant.

SURVEY RESULTS

Of the 444 survey booklets sent over the 30-month study period, 367 were returned. Surveys were sent at the end of baseline, the end of Year 1, and the end of Year 2 of each subject's study enrollment. All subjects enrolled at the time of each survey distribution are included in the analysis. The control group returned 74% of their surveys, the telephone group returned 83%, and the video group returned 89% of the surveys. Overall, 82% of all surveys were completed and returned. There were no significant differences in survey return rates between study groups.

Although 70% of the intervention subjects returned all three surveys, only 40% of the controls did so. Conversely, 11% of the control subjects and 1.5% of the intervention subjects did not return any surveys. These findings were consistent with comments reported as part of the Year 2 survey, where almost all intervention subjects (telephone and video) but fewer than a quarter of control subjects thought the study was useful.

Mean return intervals for surveys ranged from 43 days (video group) to 87 days (control group). There were no statistically significant differences in survey return intervals among study groups at each of the survey collection periods.

The effect of incentives on survey return rates is given in *Figure 4*. Return rates were 55% with no incentive and increased to 70% with the $2.00 incentive and to 95% with the $20 gift card. Incentives were added as part of a bundled effort to improve collection returns, including the addition of staff and reformatting surveys into a more accessible booklet format. Thus it was not possible to determine the unique influences of these changes on return rates.

Discussion

This article describes an observational study of data collection behaviors of families of children with medical complexity participating in the TeleFamilies RCT. High rates of return for both monthly HCSU calendars and annual health-related surveys are attainable but require a flexible and
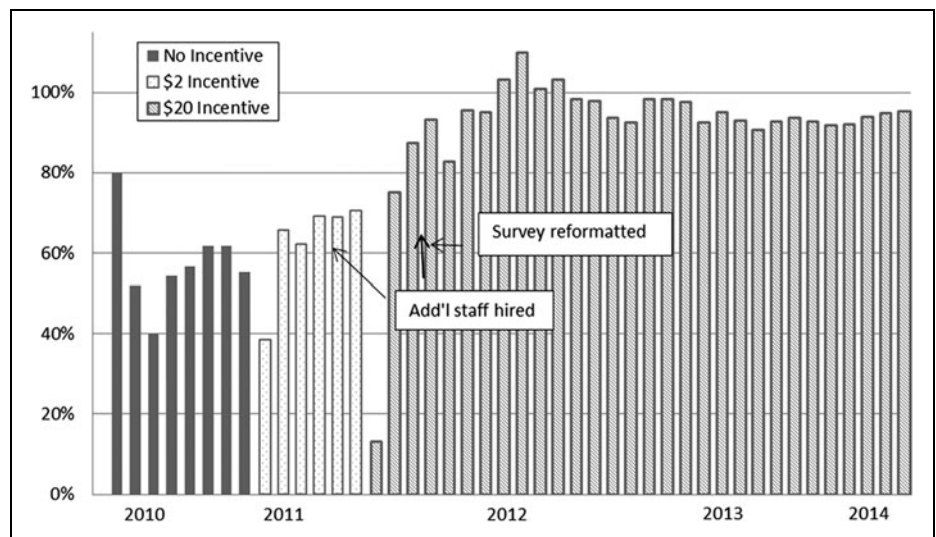

Fig. 4. Monthly survey return rates over time and incentive program. The extreme high and low response rates were due to the time lag in sending and receiving surveys in any given month.

personnel-intensive approach. Using multifaceted methods for data collection, return rates of approximately 90% for HCSU calendars and 82% for annual surveys were achieved. Lessons learned include the need to tailor efforts to family preferences and collecting this preference information at enrollment to optimize data collection. Identifying the preferred mode (mail, e-mail, telephone, in-person) for a given family was itself an experimental process that required flexibility of study procedures and research staff. Adapting incentive types was met with increased response rates over time.

A recent study using the Technology Acceptance Model identified four universal predictors for acceptance of home telehealth services: perceived usefulness, ease of use, the influence of important others, and facilitating support conditions.[22] These predictors were not explicitly driving the TeleFamilies data collection effort, but awareness of these predictors and the application of Social Exchange Theory likely facilitated exceptional response rates from this sample. The success of the TeleFamilies data collection process suggests that these universal predictors apply to other health support systems and should be considered when developing and implementing new approaches to healthcare.

EMR data are increasingly used to monitor and evaluate the quality of care. There are several significant limitations to EMR data, including variation in structured data elements across vendors, gaps in out-of-network care, and clinical workflows that are reliant on narrative documentation.[23] Administrative data, EMR data, or survey data alone are not equipped to capture accurate information on healthcare utilization for children with medical complexity.[24] Due to the high variability and "noisy" nature of these data for this population, imprecise data pose the risk of missing critical details that could guide decisions about clinical care delivery, risk adjustment, and healthcare policy. Using a multimodal approach to collecting and validating HCSU and outcome data in this study addressed many of these limitations.

Recall error is a limitation of methods that rely on respondents to report on behaviors or service use. Staff validation of HCSU calendars with both families and EMR review was a successful strategy in this study. When children have frequent healthcare encounters, caregivers may underestimate how often their child is seeing providers. Collecting these data using a calendar and staff resources to validate the data is acceptable to respondents and likely captures more accurate data than surveys that require annual HCSU recall. In another study, interviews assessed services, supports, and costs for families under stress.[9] These investigators found that parents rarely had problems answering the survey questions or recalling services used in the prior 3-month period. This supports the reliability of parental recall within a short time period.

In TeleFamilies, survey response rates declined after a successful start. The addition of study personnel, an improved survey booklet format, and financial incentives increased response rates to approximately 95%. These results contrast with several reports in the literature on the effect of incentives on response rates in healthcare studies.[7,8] The high response rate for TeleFamilies may also be a result of the perceived health and quality of life benefit of study participants that may not have been the case in the other populations. Thus, although financial incentives may have been important components in the data collection strategy, there were other factors whose impact could not be separated from the incentives as drivers of improving response patterns.

This study has implications and limitations that might affect generalizability of our findings. This study demonstrated that it is possible to attain a 90% response rate for HCSU in an institution that maintains an EMR. HCSU data obtained from the EMR could be confirmed by subject families to assure completeness and correctness of the record, and EMR records could confirm subject recall in HCSU calendar reporting. However, relying solely on the EMR for HCSU may not reflect services obtained at other institutions. Another potential limitation to generalizability is the difference in survey response rates demonstrated over the 30-month study period, during which time approximately 70% of intervention subjects returned surveys at all three collection times but only 40% of controls returned all three surveys. These findings could impact between-group comparisons over time. Finally, the observational nature of study findings related to data collection modes, incentives, staffing and survey presentations make it difficult to draw general conclusions thath might have been possible in a randomized clinical trial of data collection methods, which was not a part of the overarching TeleFamilies study.

Acknowledgments

Disclosure Statement

No competing financial interests exist.

REFERENCES

1. Groves R, Lyberg L Total survey error: Past, present, and future. *Public Opin Q* **2010;**74:849–879.

2. Meinert CL. Clinical trials and treatment effects monitoring. *Control Clin Trials* **1998;**19:515–522.

3. Dixon BE, Grannis SJ, Revere D. Measuring the impact of a health information exchange intervention on provider-based notifiable disease reporting using mixed methods: A study protocol. *BMC Med Inform Decis Mak* **2013;**13:121–129.

4. Larsen IK, Smastuen M, Johannesen TB, Langmark F, Parkin DM, Bray F, Moller B. Data quality at the Cancer Registry of Norway: An overview of comparability, completeness, validity, and timeliness *Eur J Cancer* **2009;**45:1218–1231.

5. Dillman DA. *Mail and Internet surveys—The tailored design method*, 2nd ed. Hoboken, NJ: John Wiley & Sons, **2007.**

6. Dillman DA, Phelps G, Tortora R, Swift K, Kohrell J, Berck J, Messer BL. Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet. *Soc Sci Res* **2009;**38:1–18.

7. Nakash RA, Hutton JL, Jorstad-Stein EC, Gates S, Lamb SE. Maximising response to postal questionnaires—A systematic review of randomized trials in health research. *BMC Med Res Methodol* **2006;**6:5–14.

8. Glidewell L, Thomas R, MacLennan G, Bonetti D, Johnston M, Eccles MP, Edlin R, Pitts NB, Clarkson J, Steen N, Grimshaw JM. Do incentives, reminders or reduced burden improve healthcare professional response rates in postal questionnaires? Two randomized controlled trials. *BMC Health Serv Res* **2012;**12:250–259.

9. Sleed M, Beecham J, Knapp M, McAuley C, McCurry N. Assessing services, supports, and costs for young families under stress. *Child Care Health Dev* **2005;**32:101–110.

10. Cropanzano R, Mitchell MS. Social exchange theory: An interdisciplinary review. *J Manage* **2005;**31:874–900.

11. Looman WS, Presler E, Erickson MM, Garwick AW, Cady RG, Finkelstein SM. Care coordination for children with complex special healthcare needs: The value of the advanced practice nurse's enhanced scope of knowledge and practice. *J Pediatr Health Care* **2013;**27:293–303.

12. Cohen E, Kuo D, Agrawal R, Berry J, Bhagat S, Simon T, et al. Children with medical complexity: An emerging population for clinical and research initiatives. *Pediatrics* **2011;**127:529–538.

13. American Academy of Pediatrics. National Center for Medical Home Implementation. **2013.** Available at http://www.medicalhomeinfo.org (last accessed February 11, 2014).

14. Bramlett MD, Read D, Bethell C, Blumberg SJ. Differentiating subgroups of children with special healthcare needs by health status and complexity of healthcare needs. *Matern Child Health J* **2009;**13:151–163.

15. Cady RG, Erickson M, Lunos S, Finkelstein SM, Looman W, Celebrezze M, Garwick A. Meeting the needs of children with medical complexity using a telehealth advanced practice registered nurse care coordination model. *Matern Child Health J* **2015;**19:1497–1506.

16. Varni JW, Burwinkle TM, Seid M, Skarr D. The PedsQL 4.0 as a pediatric population health measure: Feasibility, reliability, and validity. *Ambul Pediatr* **2003;**3:329–341.

17. Varni JW, Sherman SA, Burwinkle TM, Dickinson PE, Dixon P. The PedsQL family impact module: Preliminary reliability and validity. *Health Qual Life Outcomes* **2004;**2:55–61.

18. Crofton C, Darby C, Farquhar M, Clancy CM. The CAHPS hospital survey: Development, testing, and use. *Jt Comm J Qual Patient Saf* **2005;**31:655–659.

19. Looman WS. Development and testing of the social capital scale for families of children with special healthcare needs *Res Nurs Health* **2006;**29:325–336.

20. Stein RE, Jessop DJ. Functional Status II®. A measure of child health status. *Med Care* **1990;**28:1041–1055.

21. Demiris G, Speedie S, Finkelstein S. A questionnaire for the assessment of patients' impressions of the risks and benefits of home telecare. *J Telemed Telecare* **2000;**6:278–284.

22. Cimperman M, Brenčič M, Trkman P, de Leonni Stanonik M. Older adults' perceptions of home telehealth services. *Telemed J E Health* **2013;**19:786–790.

23. Bailey LC, Mistry KB, Tinoco A, Earls M, Rallins MC, Hanley K, Woods D. Addressing electronic clinical information in the construction of quality measures. *Acad Pediatr* **2014;**14(5 Suppl):S82–S89.

24. Berry JG, Agrawal R, Kuo DZ, Cohen E, Risko W, Hall M, Casey P, Gordon J, Srivastava R. Characteristics of hospitalizations for patients who use a structured clinical care program for children with medical complexity. *J Pediatr* **2011;**159:284–290.

Address correspondence to:
Stanley M. Finkelstein, PhD
Department of Laboratory Medicine
and Pathology/Health Informatics
MMC 609
University of Minnesota
Minneapolis, MN 55455

E-mail: stan@umn.edu