Listening to the nurse pays off: an integrated Nurse HealthLine programme was associated with significant cost savings

JESSICA L. NAVRATIL-STRAWN MS[1], KEVIN HAWKINS PhD[2], TIMOTHY S. WELLS MPH, PhD[3], RONALD J. OZMINKOWSKI PhD[4], JEAN HAWKINS-KOCH BBA[5], HUNGCHING CHAN MPH[6], STEPHEN K. HARTLEY BS[7], RICHARD J. MIGLIORI MD[8] and CHARLOTTE S. YEH MD[9]

[1]Senior Research Analyst, *OptumHealth, Golden Valley, MN,* [2]Senior Research Director, [3]Senior Health Economics Researcher, *Advanced Analytics, OptumInsight, Ann Arbor, MI,* [4]Chief Scientific Officer and Senior Vice President, *Health-care Analytics, OptumHealth Care Solutions, Ann Arbor, MI,* [5]Senior Market Director, *Public Sector, OptumHealth,* [6]Vice President, *Clinical Analytics, OptumHealth, Golden Valley, MN,* [7]Senior Director of Analytics, *Health and Wellness, OptumHealth, Phoenix, AZ,* [8]Executive Vice President, *Health Services, UnitedHealth Group, Minnetonka, MN and* [9]Chief Medical Officer, *AARP Services, Inc., Washington, DC, USA*

Correspondence
Timothy S. Wells
Advanced Analytics
OptumInsight
315. E. Eisenhower Parkway,
Suite 305
Ann Arbor
MI 48108
USA
E-mail: tim.wells@optum.com

NAVRATIL-STRAWN J.L., HAWKINS K., WELLS T.S., OZMINKOWSKI R.J., HAWKINS-KOCH J., CHAN H., HARTLEY S.K., MIGLIORI R.J. & YEH C.S. (2014) *Journal of Nursing Management* **22**, 837–847.

Listening to the nurse pays off: an integrated Nurse HealthLine programme was associated with significant cost savings

Aim To estimate the relationship between adherence to nurse recommendations about where to seek care and expenditures for health-care services received by callers to a Nurse HealthLine telephone-based triage programme.
Methods Health-care utilization and claims data from callers to the Nurse HealthLine were included. Adherent callers were those who followed the nurse recommendations, while those who did not were classified as non-adherent. Programme-related savings were estimated using differences in downstream health-care expenditures between adherent and non-adherent callers after using multivariate modelling to adjust for between-group differences.
Results Fifty-five per cent of callers were adherent. Nurses were over three times as likely (41% *vs.* 13%) to recommend seeking a higher level of care (e.g. emergency room *vs.* urgent care). Regression analyses showed that the impact of getting members to the appropriate place of care was associated with significant annual savings of $13.8 million ($P < 0.05$), attributable mostly to Medicare, generating a positive return on investment of $1.59.
Conclusions This is the first known Nurse HealthLine triage programme exclusively for Medicare beneficiaries with supplemental coverage.
Implications for nursing management Nurse managers should consider promoting telephone-based triage programmes as complementary to clinical nursing, which has a direct impact on health-care utilization and costs.

Keywords: economics, Medicare, nurse, telephone hotline

Accepted for publication: 30 November 2012

Background

A rapidly growing area of clinical practice is the offering of advice over the telephone about where to seek treatment (Rutenberg 2000, Valanis *et al.* 2003). Such telephone-based triage programmes allow an individual to speak with a registered nurse about health-related activities, including current health concerns.

DOI: 10.1111/jonm.12048

Based upon the caller's symptoms, the nurse will assess the urgency of the caller's need for care and make a recommendation about where to seek health-care (Robinson *et al.* 1996, Sabin 1998, Cariello 2003).

When these programmes were first offered, there was discomfort over the acceptance and safety of telephone triage among clinicians and administrators alike (Robinson *et al.* 1997, Allen-Davis *et al.* 2002, Hildebrandt *et al.* 2003, Kempe *et al.* 2006). To help alleviate these concerns, many triage centers developed programmes that included registered nurses with specialized training and access to computer-based clinical protocols and algorithms (Sabin 1998, Melzer & Poole 1999, Nauright *et al.* 1999, Cariello 2003). Studies now show that callers routinely receive appropriate health-care management advice, that telephone triage does not typically delay or ration care and that satisfaction with triage programmes is high (O'Connell *et al.* 2001b, Moore *et al.* 2002, Lee *et al.* 2003, Dent 2010) These findings are reassuring, as about 100 million US citizens had access to telephone triage systems as of 2001 (Lancet 2001).

From a programme management standpoint, an important key to the success of telephone triage systems is their ability to be cost effective, or at least cost neutral while providing improved quality or added services. Several studies suggest that telephone triage frequently refers individuals to an appropriate but lesser level of care than otherwise would have been sought, which has reduced associated costs (Wolf-Klein & Silverstone 1987, Barber *et al.* 2000, Frisbee *et al.* 2001). However, return on investment (ROI) analyses of these programmes are rarely provided. One study compared average health-care expenditures in the year following implementing a telephone triage system in a health maintenance organisation (HMO) and point-of-service insured population, and estimated overall savings of $1.70 per dollar spent on the programme (i.e. a ROI ratio of 1.70) (O'Connell *et al.* 2001a). Other telephone-based triage programmes published in the literature that reported savings calculations involved paediatric populations (Cariello 2003, Bunik *et al.* 2007) or an urban safety-net health network (Bogdan *et al.* 2004) but did not provide information on ROI. This leaves a gap in the knowledge regarding the use of telephone triage, especially in older populations. The primary objective of this study is to estimate the ROI associated with the Nurse HealthLine triage service.

Methods

Intervention

The Nurse HealthLine service is available to over 3.1 million individuals who have purchased an AARP Medicare Supplement Insurance plan provided by UnitedHealthcare Insurance Company (for New York residents, UnitedHealthcare Insurance Company of New York, NY, USA). These plans are offered in all 50 states, Washington, DC, and various territories of the USA. The Nurse HealthLine is a 24-hour service that provides two main services, triage and non-triage. Triage calls comprise about 46% of the call volume. These calls involve those in which Nurse HealthLine staff suggests the most appropriate place to seek health-care for those calling about acute physical or mental health problems. During triage calls, a nurse recommends a particular site of care consistent with member symptoms. Sorted by level of intensity, this may be an emergency room, an urgent care center, a visit to a doctor's office or self-treatment at home.

Before making a recommendation about where to seek care, the nurse also asks the caller where he or she was intending to seek care. The nurse may then either confirm the caller's original intent or recommend a more or less intensive level of care. Non-triage calls account for 54% of the call volume. These involve calls about insurance benefit questions and calls in which Nurse HealthLine staff assist with referrals to community services and health-care providers. In order to work effectively in this position, Nurse HealthLine nurses are all registered nurses with AARP certification and Medicare immersion training. Certification requires each nurse to go through an instructor-led and internally developed programme designed to cover multiple aspects of the Medicare population and the Nurse HealthLine programme. Training includes a variety of activities such as listening to recorded calls in which the personal information has been removed, role play, scenarios and problem solving, as well as several computer based training modules. At the end of the training, nurses must pass a post-training examination to become certified.

Study sample

This study is based upon information from Nurse HealthLine triage calls from insured members placed between 1 January 2010 and 31 December 2010. To evaluate the impact of nurse-led conversations, the authors defined a pre-call period as the 6 months before the call and a post-call observation period as

the 30 days after the call. Although callers were under no obligation to follow nurse recommendations, callers who followed nurse recommendations regarding an appropriate place to seek health-care were classified as adherent callers, while callers who did not follow nurse recommendations were classified as non-adherent callers for the purposes of this study.

The evaluation was conducted using the call as the level of observation. Some members made multiple calls throughout the study period; only calls that were at least 30 days apart were included in this study. Subsequent calls <30 days after the initial call were excluded to assess adherence to the nurse's original recommendation regarding where to seek care more easily. Other exclusions involved calls by members who were not continuously enrolled in an AARP Medicare Supplement Insurance plan during the pre-call and post-call observation periods. Calls from members with missing demographic or risk score information were excluded. For each care path, call records associated with members whose expenditures were above the 98th percentile (i.e. generally >$65 000, but up to a maximum of $996 000) were removed from analysis, as were records for those with negative expenditures in the pre-call or post-call periods. These records were excluded to avoid undue influence of calls associated with outlier expenditures.

Care paths

Study groups and outcomes for this analysis were determined by tracking the movement of each caller between pre-call intention, nurse recommendations and actual services used, based on analyses of claims data. The spectrum of care ranges from the highest (emergency room) and lowest level (self care) and was ranked as follows: emergency room > urgent care > medical doctor > self care. The movement in care levels from pre-intent to claims validated post-intent was categorized in three ways, which yielded five possible care paths:

- From emergency room to a lower level of care: the caller's pre-call intent was to seek care in an emergency room but the nurse recommended a lower level of care, such as an urgent care visit or a visit to a doctor's office or self-treatment at home.
- From non-emergency room/office visit to a lower level of care: the caller's pre-intent was to seek non-emergency care via an urgent care visit or visit to a doctor's office, but the nurse recommended a lower level of service.
- Same level of care: the nurse agreed with the caller's pre-call intent and then provided information to prepare the member for that level of care.
- From non-emergency room level of care to emergency room: the caller's pre-call intent was not to go to the emergency room but the nurse recommended a visit to an emergency room.
- From non-emergency room visit to higher, non-emergency room level of care: the nurse recommended a more intensive site of care than the caller was originally intending, but not care in an emergency room.

Adherence to the recommended care path

Determining adherence to nurse recommendations was measured by reviewing claims data for up to 30 days after the call was made. Adherence to a recommendation of emergency room treatment was defined based on whether such a visit was made within 2 days of the call. Adherence to a recommended visit to a doctor's office was determined based on whether the visit occurred within 14 days of the call. Adherence to a recommendation of self-care was evident if no emergency room visit occurred within 2 days of the call and if no doctor's office visit occurred within 14 days of the call.

There were two limitations in the data that affected the assignment of whether or not callers followed nurse recommendations. These included instances when nurse recommendations were to seek treatment at an urgent care clinic and when nurse recommendations were to call their doctor. First, it was not always possible to identify urgent care claims in these data, and second, claims data could not be used to validate a call to a doctor. To adjust for these limitations, callers were considered adherent to a recommendation to go to urgent care if claims data showed a visit to an urgent care clinic or doctor's office within 2 days of the call. If the nurse recommendation was to call their doctor, the caller was considered adherent if the claims data showed no claim within 14 days, or a claim for a visit to a doctor's office within 14 days of the call. Expenditures associated with each care path were then calculated and categorized regarding adherence based on whether the caller used the nurse-recommended site of care.

Return on investment estimation

The ROI associated with adherence to the nurse recommendations was estimated as a ratio of Nurse HealthLine programme savings divided by its costs to

AARP. An ROI ratio >1.0 implies that for every dollar invested by AARP, more than a dollar was returned in savings by following the nurse recommendations.

Savings were estimated as the difference in average post-call period expenditures for adherent callers *vs.* post-call expenditures for non-adherent callers. The difference in the average amount saved by adhering was then multiplied by the number who were adherent, to estimate savings associated with each care path. Analyses were done separately for each care path because caller severity/urgency varied by care path. Savings or losses from all five care paths were then added to produce an overall savings or loss estimate for the Nurse HealthLine programme.

Variable creation

Demographic measures included the caller's age, gender and two variables measuring location. The first indicated whether the caller resided in a rural or an urban location. The second variable indicated the census region of the country the caller resided in (Northeast, Midwest, South and West).

Socioeconomic variables included imputations of the caller's race and income. These were geocoded based on the zip code where the caller resided. Race was categorized as high, medium or low, depending on the percentage of minority residents in the caller's zip code. Income was inferred as high, medium-high, medium or low based on whether the median income in the caller's zip code area was in the highest, second-highest, third-highest or lowest quartile in 2010, according to US Census records.

Pharmaceutical claims data were available for approximately half of those included in these analyses (i.e. those who had Medicare Part D coverage provided by UnitedHealthcare Insurance Company). A binary variable was created to account for the impact of having pharmacy expenditure in the 6 months before the Nurse HealthLine call.

Several health status measures were also included in the analyses. First among these was the OptumInsight ImpactPro prospective risk score (http://www.optuminsight.com/content/attachments/ImpactProforCareManagement.pdf), which predicted future Medicare-allowed charges. The remaining health status variables included health-care utilization in the 6 months before the call.

The analyses also accounted for differences in the supply of health-care services in the areas where callers lived, because these are known to influence health-care utilization and expenditures (Wennberg *et al.* 2002). We therefore included measures based on the number of primary care physicians, specialists and hospital beds in the caller's zip code of residence. The number of physicians and specialists were calculated per 100 000 residents, while the number of hospital beds was calculated per 1000 residents in the caller's hospital service area (Wennberg & Cooper 1999).

We also accounted for the reason why the caller called the Nurse HealthLine. The caller's chief complaint was categorized as follows: circulatory, digestive, endocrine, genitourinary, infections, injury/poisoning, mental disorders, musculoskeletal, nervous system, respiratory, skin/subcutaneous conditions or other conditions. Next, we categorized calls as occurring during a weekday or weekend, as weekend calls may be more urgent. The nurse's recommended place of service (e.g. emergency room) was also captured. Finally, we accounted for whether the call included in the analysis was from a member who called Nurse HealthLine only once in 2010 *vs.* from members who called two or more times during 2010.

Analyses

Two sets of analyses were performed. The first set was unweighted and was used to illustrate baseline case-mix differences between adherent and non-adherent groups. Chi-square and Student's *t*-tests were used to test for statistical differences in the categorical and continuous variables between the two groups.

The second set of analyses used propensity score weighting (Rosenbaum & Rubin 1983, Rosenbaum 1987, Morgan & Todd 2008, Austin 2011) to minimize case-mix differences between the adherent and non-adherent groups.

Within each care path, a logistic regression analysis estimated the likelihood of being an adherent caller. Based on the logistic regression results, a propensity score for each caller was defined as his or her predicted probability of being adherent to nurse recommendations. This probability was used to construct a case weight, which was used to re-estimate the independent and health-care expenditure variables.

Once the propensity score analyses were completed, an examination of the differences between weighted independent variables between the adherent and non-adherent callers was conducted to determine if the propensity score weights adequately accounted for case-mix differences. If there were no case-mix differences between the two groups after weighting, then any differences in the weighted health-care expenditures were likely to be attributable to following the nurse's advice. Savings occurred when weighted average expenditures

for adherent group members were less than weighted average expenditures for non-adherent group members, while losses occurred when weighted average expenditures for adherent group members exceeded weighted averages for non-adherent group members.

Sensitivity analyses

Two sensitivity analyses were performed. The first analysis adjusted for the length of follow-up. A 30-day follow-up may be insufficient to capture the full impact of the Nurse HealthLine call, therefore the analysis was repeated using a 60-day follow-up.

The second analysis determined the impact of outliers (group members with extremely high or low expenditures in pre- or post-periods). This sensitivity analysis re-estimated savings and losses after including members with outlier expenditures. Outliers were determined separately for each care path, using a method first described by Heckman *et al.* (1997) The intention is to ensure that the ranges of health-care expenditures from lowest to highest were comparable for adherent and non-adherent callers in each care path. Expenditures outside the common ranges in each period were labelled as outliers and removed from the main analyses, but included in the sensitivity analyses.

Results

The starting sample consisted of 63 277 calls. Of these 15.9% (10 071) were excluded from the main analyses. Following exclusions for reasons described earlier, the main analyses were based on 53 206 calls made by 29 438 adherent and 23 768 non-adherent callers.

For about 45% of the calls, the nurses agreed with the pre-intent of the caller and for the other 55% recommendations differed. A lower level of care was recommended 13.4% of the time by the nurses, and these recommendations were about equally likely to be for 'from emergency room to a lower level of care' (6.6%) or 'from non-emergency room/office visit to a lower level of care' (6.8%). Conversely, 41.4% of the time the recommendation was for a higher level of care. Recommendations for 'from non-emergency room level of care to emergency room' were more frequent than were recommendations for 'from non-emergency room visit to higher, non-emergency room level of care' (36.0 and 5.4%, respectively).

Tables 1–3 show the unweighted descriptive statistics for adherent and non-adherent callers in each care path. Generally, before weighting, there were significant differences between adherent and non-adherent callers for most variables. After weighting, all significant differences were removed between adherent and non-adherent callers. This implies that the propensity-score weighting removed measurable case mix differences, allowing the impact of the program to be estimated using the weighted post-call expenditure values.

Average 30-day medical and prescription expenditures were calculated for each care path (Table 4). The weighted results show significant savings for the programme per caller per month for four of the five care paths. Specifically, 'from emergency room to a lower level of care' had the greatest savings ($4688), followed by 'from non-emergency room/office visit to a lower level of care' ($1381), 'from non-emergency room visit to higher, non-emergency room level of care' ($453) and 'same level of care' ($330), when comparing the difference between adherent and non-adherent callers. As expected, 'from non-emergency room level of care to emergency room' expenditures for adherent callers were $6057 higher in comparison with the non-adherent callers.

Savings per adherent caller and total savings are shown in Table 5. Total savings, calculated as the average savings per adherent caller multiplied by the number of adherent callers for each care path, and then summed across all care paths, were estimated at about $13.8 million ($P < 0.01$) for the Nurse HealthLine. Among the four care paths that had savings, 48.8% were attributable to 'from emergency room to a lower level of care', 11.4% to 'from non-emergency room/office visit to a lower level of care', 20.2% to 'same level of care' and 19.6% to 'from non-emergency room visit to higher, non-emergency room level of care'. When nurse recommendations were from non-emergency room level of care to emergency room there was a loss of over $8.9 million.

Total programme costs were required to calculate the programme's ROI. These costs were determined by first totalling all of the invoices paid for programme services, which amounted to $8.7 million. Subsequently, the ROI was determined by dividing the estimated programme savings of over $13.8 million by the cost of the programme. As a result, we estimated an ROI of 1.59 ($P < 0.01$) for the programme. This is similar to those reported by O'Connell *et al.* (2001a) This implies that for every dollar spent on the programme, $1.59 was saved. The savings are mostly a result of reduced downstream medical and prescription pharmacy expenditures.

Sensitivity analyses results

The detailed results of the sensitivity analyses are not presented here, but are available upon request. Extend-

Table 1
Unweighted descriptive statistics for the 'lower level of care' groups

	From emergency room			*From non-emergency room/office visit*		
Variable	*Adherent caller (n = 2376)*	*Non-adherent caller (n = 1173)*	*P*	*Adherent caller (n = 1884)*	*Non-adherent caller (n = 1768)*	*P*
Average age (years)	76.9	77.2	0.184	76.8	77.6	0.004
Sex (% female)	65.1	64.7	0.799	71.3	71.8	0.735
Metropolitan statistical area (% MSA)	78.6	77.5	0.408	81.1	83.8	0.028
Previous 6-month emergency room visit count*	0.6	0.7	0.044	0.3	0.6	<0.001
Previous 6-month inpatient admission count*	0.1	0.2	0.090	0.1	0.1	<0.001
Previous 6-month physician's office visit count*	6.9	7.3	0.014	5.2	8.1	<0.001
Prospective 3-month cost risk score†	6.2	8.2	<0.001	4.0	6.2	<0.001
Hospital beds (per 1000)‡	2.3	2.3	0.310	2.3	2.3	0.094
Primary care physicians (per 100 000)‡	65.4	64.6	0.444	65.2	67.4	0.020
Specialists (per 100 000)‡	122.3	123.4	0.527	123.6	126.6	0.064
Nurse HealthLine chief complaint (%)						
Circulatory system	14.1	13.5	0.035	5.3	5.0	0.004
Digestive system	10.5	10.1		10.2	11.4	
Endocrine system	3.3	4.1		2.7	3.6	
Genitourinary system	4.8	4.7		4.2	3.5	
Infectious	1.4	1.5		2.7	2.8	
Injury and poisoning	14.4	18.5		22.2	20.7	
Mental disorder	1.2	1.1		1.3	1.8	
Musculoskeletal system	12.0	9.4		13.8	11.0	
Nervous system	11.7	9.7		10.0	9.5	
Respiratory system	8.3	7.7		11.4	11.7	
Skin and subcutaneous tissue	3.2	3.0		7.2	6.2	
Other	15.2	16.8		9.1	12.8	
Day of Nurse HealthLine call (% weekend)	40.0	44.0	0.021	30.8	27.1	0.011
Nurse recommendations (%)						
Urgent care	1.8	11.7	<0.001	27.7	3.8	<0.001
Visit to physician's office	86.8	55.6		72.3	96.2	
Self-care	11.3	32.8		2.1	5.1	
Repeat caller (%)	3.4	2.6	0.182	56.5	58.0	0.336
Prescription benefit use (%)§	56.5	58.0	0.336	27.7	3.8	<0.001
Income (%)‡						
High	46.8	44.4	0.279	49.2	50.2	0.703
Upper medium	23.9	26.9		24.2	24.9	
Lower medium	20.2	19.8		18.5	17.0	
Low	9.1	8.9		8.0	7.9	
Minority status (%)‡						
Low	58.0	58.6	0.108	58.6	56.7	0.278
Medium	37.9	38.8		38.6	39.8	
High	4.0	2.6		2.8	3.6	
Region (%)‡						
Northeast	19.5	21.6	0.552	22.7	27.1	<0.001
Midwest	17.4	17.2		18.8	15.3	
South	36.6	35.9		32.9	35.4	
West	26.5	25.3		25.7	22.2	

MSA, metropolitan statistical area.
*Expressed as a percentage based upon per person counts.
†Based upon OptumInsight ImpactPro prospective risk scores.
‡Based upon caller's zip code.
§Based upon per cent who used a Part B prescription drug plan.

ing the follow-up period to 60 days yielded a slightly lower ROI ratio of $1.44 in savings per dollar spent on the programme. The second sensitivity analysis found that the results were sensitive to a few callers with either very high or very low expenditures in the pre- or post-periods. When about 0.7% ($n = 456$) of

Table 2
Unweighted descriptive statistics for the 'same level of care' group

Variable	*Adherent caller (n = 13 991)*	*Non-adherent caller (n = 10 262)*	*P*
Average age (years)	77.5	77.8	0.004
Sex (% female)	70.4	69.6	0.152
Metropolitan statistical area (% MSA)	80.3	80.1	0.686
Previous 6-month emergency room visit count*	0.4	0.5	<0.001
Previous 6-month inpatient admission count*	0.1	0.1	0.001
Previous 6-month physician's office visit count*	6.3	6.7	<0.001
Prospective 3-month cost risk score†	5.3	5.9	<0.001
Hospital beds (per 1000)‡	2.3	2.3	0.002
Primary care physicians (per 100 000)‡	66.0	67.0	0.011
Specialists (per 100 000)‡	124.3	125.0	0.299
Nurse HealthLine chief complaint (%)			
Circulatory system	9.4	7.6	<0.001
Digestive system	11.8	13.5	
Endocrine system	2.1	2.8	
Genitourinary system	4.9	4.2	
Infectious	2.2	1.8	
Injury and poisoning	16.8	18.0	
Mental disorder	1.8	2.1	
Musculoskeletal system	13.6	12.9	
Nervous system	9.7	9.7	
Respiratory system	9.3	9.9	
Skin and subcutaneous tissue	6.4	4.9	
Other	12.0	12.6	
Day of Nurse HealthLine call (% weekend)	27.5	29.3	0.002
Nurse recommendations (%)			
Emergency room	7.8	4.9	<0.001
Urgent care	0.3	0.4	
Visit to physician's office	55.0	47.2	
Self-care	36.9	47.5	
Repeat caller (%)	9.7	10.5	0.019
Prescription expense (%)§	57.7	57.6	0.843
Income (%)‡			
High	48.0	48.2	0.413
Upper medium	25.6	24.9	
Lower medium	17.8	18.4	
Low	8.6	8.6	
Minority status (%)‡			
Low	58.6	58.4	0.906
Medium	38.1	38.3	
High	3.4	3.3	
Region (%)‡			
Northeast	24.0	25.2	0.024
Midwest	17.2	18.0	
South	35.9	34.6	
West	22.8	22.1	

MSA, metropolitan statistical area.
*Expressed as a percentage based upon per person counts.
†Based upon OptumInsight ImpactPro prospective risk scores.
‡Based upon caller's zip code.
§Based upon per cent who used a Part B prescription drug plan.

the calls considered to be outliers were added back into the analysis, the total savings were $22.1 million, and the ROI increased to $2.53 per dollar spent on the programme. The inclusion, or omission of outliers is often of interest to programme evaluation stakeholders. Many who are conversant on this subject argue for or against the inclusion of outliers. We prefer to include both results, with and without outliers, to provide the greatest context for the reader.

Discussion

To our knowledge, this study is unique in two ways. It may be the first evaluation of a telephone-based

Table 3
Unweighted descriptive statistics for the 'higher level of care' groups

	To emergency room			*To non-emergency room/office visit*		
Variable	*Adherent caller (n = 1501)*	*Non-adherent caller (n = 1407)*	*P*	*Adherent caller (n = 9918)*	*Non-adherent caller (n = 9382)*	*P*
Average age (years)	77.3	77.7	0.177	78.0	78.2	0.085
Sex (% female)	67.8	67.3	0.756	69.1	69.1	0.942
Metropolitan statistical area (% MSA)	79.4	76.9	0.088	80.8	78.3	<0.001
Previous 6-month emergency room visit count*	0.6	0.6	0.706	0.5	0.4	<0.001
Previous 6-month inpatient admission count*	0.2	0.2	0.279	0.1	0.1	<0.001
Previous 6-month physician's office visit count*	6.7	7.2	0.009	7.7	5.3	<0.001
Prospective 3-month cost risk score†	8.0	7.1	0.001	5.3	5.9	<0.001
Hospital beds (per 1000)‡	2.3	2.3	0.946	2.3	2.3	<0.001
Primary care physicians (per 100 000)‡	64.1	67.1	0.004	66.0	66.6	0.156
Specialists (per 100 000)‡	120.2	122.1	0.274	123.3	123.6	0.657
Nurse HealthLine chief complaint (%)						
Circulatory system	34.8	30.1	<0.001	9.3	8.3	<0.001
Digestive system	5.8	3.6		13.9	16.2	
Endocrine system	2.1	3.5		3.2	3.1	
Genitourinary system	3.9	3.4		4.7	4.1	
Infectious	0.6	0.3		2.5	1.8	
Injury and poisoning	12.2	14.5		11.6	11.7	
Mental disorder	1.2	1.9		2.4	2.5	
Musculoskeletal system	7.0	8.3		14.9	14.6	
Nervous system	9.0	11.7		11.4	11.5	
Respiratory system	6.0	6.2		8.1	7.4	
Skin and subcutaneous tissue	0.4	0.5		6.1	6.0	
Other	17.0	16.1		11.9	12.8	
Day of Nurse HealthLine call (% weekend)	39.8	34.0	0.001	31.8	33.1	0.040
Repeat caller (%)	2.1	3.2	0.045	9.1	8.4	0.083
Prescription expense (%)§	58.9	59.0	0.985	58.2	55.5	<0.001
Income (%)‡						
High	47.0	44.3	0.431	48.2	45.5	0.001
Upper medium	25.7	26.2		25.1	25.9	
Lower medium	19.3	21.4		18.0	19.8	
Low	8.0	8.1		8.8	8.8	
Minority status (%)‡						
Low	62.3	57.9	0.013	59.0	60.1	0.906
Medium	35.3	38.2		38.1	36.6	
High	2.5	3.9		2.9	3.2	
Region (%)‡						
Northeast	16.6	18.5	0.195	23.2	22.9	0.024
Midwest	20.3	17.5		16.8	19.6	
South	36.7	37.4		35.9	35.1	
West	26.3	26.7		24.1	22.4	
Nurse recommendations (%)						
Urgent care	n/a	n/a		2.4	5.4	<0.001
Visit to physician's office	n/a	n/a		97.6	94.6	

MSA, metropolitan statistical area.
*Expressed as a percentage based upon per person counts.
†Based upon OptumInsight ImpactPro prospective risk scores.
‡Based upon caller's zip code.
§Based upon per cent who used a Part B prescription drug plan.

Table 4
Average 30-day medical and prescription expenditures by care path

	Unweighted				*Weighted**			
Care path	*Adherent caller*	*Non-adherent caller*	*Difference†*	*P*	*Adherent caller*	*Non-adherent caller*	*Difference†*	*P*
From emergency room to a lower level of care	$2636	$7385	($4749)	<0.001	$2947	$7635	($4688)	<0.001
From non-emergency room/office visit to a lower level of care	$1321	$2364	($1043)	<0.001	$1555	$2936	($1381)	<0.001
Same level of care	$2377	$2702	($325)	<0.001	$2272	$2602	($330)	<0.001
From non-emergency room visit to higher, non-emergency room level of care	$2474	$2401	$72	0.444	$2314	$2767	($453)	<0.001
From non-emergency room level of care to emergency room	$10 761	$3918	$6843	<0.001	$10 336	$4279	$6057	<0.001

*Weights based upon propensity score weighting.
†Difference between adherent and non-adherent callers indicates Nurse HealthLine programme savings, while a positive difference indicates programme losses.

Table 5
Nurse HealthLine savings and losses calculations by care path

Care Path	*Number of non-adherent callers (A)*	*Number of adherent callers (B)*	*Adherence% (C = B/(A + B))*	*Average health-care payments for adherent callers (D)*	*Average health-care payments for non-adherent callers (E)*	*Savings or losses per adherent caller (F = E – D)*	*Total savings or losses (G = B × F)*
From emergency room to a lower level of care	1151	2362	67.2%	$2947	$7635	$4688	$11 073 528
From non-emergency room/office visit to a lower level of care	1751	1871	51.7%	$1555	$2936	$1381	$2 583 851
Same level of care	10 174	13 879	57.7%	$2272	$2602	$330	$4 585 622
From non-emergency room visit to higher, non-emergency room level of care	9292	9851	51.5%	$2314	$2767	$453	$4 460 533
From non-emergency room level of care to emergency room	1400	1475	51.3%	$10 336	$4279	($6057)	($8 934 223)
Total	23 768	29 438	55.3%	$2677	$3141	$468	$13 769 311

triage programme for Medicare members with Medigap coverage, and it estimated the value of that programme by considering adherence to the nurse recommendations, which are the hallmarks of all such programmes. Among those who utilized triage services, 41% were redirected by nurses to a higher level of care, 45% were prepared for services at the same level as originally intended and 14% were recommended by nurses to seek a lower level of care. These findings imply that nurse recommendations were based on clinical evidence rather than cost considerations, as indicated by referring three times as many callers to higher rather than lower levels of care. It was also observed that nurse recommendations were well-received, as callers generally followed their recommendations, ranging between 51.3% for the 'from non-emergency room visit to higher, non-emergency room level of care' group to 67.2% for the 'from emergency room to a lower level of care' group.

Based on this evaluation, the programme resulted in significant cost savings attributable to Medicare, the Medigap insurance provider and the callers themselves, which were proportional to the costs incurred by each. About 91% of Nurse HealthLine savings were attributable to Medicare, 8% to Medigap and 1% to callers. The greatest savings were to Medicare, equating to almost $12.5 million dollars in savings during the 2010 calendar year alone. Currently, Medi-

care spending is growing at a rate about twice as fast as the US economy, placing a heavy burden on the Federal budget (Antos 2012). With budget cuts to Medicare being heavily debated as one means to reduce the US Federal expenditures (May 2011), instituting the universal availability of similar telephone-based triage programmes to all those insured with Medicare may provide appreciable cost savings in these austere times.

In addition to the data limitations surrounding urgent care visits and a call to the doctor's office described in the methods, two other limitations should be considered when interpreting the study findings. Ideally, the study would have used another Medigap population without access to a nurse health line as a comparison group to accurately measure the programme's impact; unfortunately, such a population was not available, so comparing adherent to non-adherent callers was the next best option. Analyses included total Nurse HealthLine costs (triage and non-triage), but only captured benefits from triage services in the ROI estimate. This is because it was not clear how to delineate the programme costs on a triage or non-triage basis and the author's preference to make a more conservative savings estimate. As a result, the true ROI ratio is likely to be underestimated, as any benefits for the greater than half of those calls into the program that used non-triage services were not captured.

Despite these limitations, this study had considerable strengths. Analyses were adjusted for a large number of potentially confounding variables. Meanwhile, the analytical design allowed the ROI to be calculated based upon those who followed nurse recommendations compared with those who did not. As nurse recommendation is the key feature of all triage programmes, it is important to understand the implications of compliance with nurse advice. Finally, the sensitivity analysis found that ROI estimates were relatively stable with respect to very high and very low cost outliers.

The success of this study also reflects positively on the widening role that nurses play in providing health-care. Many view the traditional role of nursing to be the provision of health-care while working closely with physicians. A key aspect for nurses in telephone triage is that it allows them to explore alternatives for treatment that will benefit the caller. This expanded role fits well with the American Nursing Association's definition of nursing that not only includes traditional clinical roles, but also reflects a broadening scope through the advocacy of care as well as interpreting patient information and making decisions about actions needed. This expanded role aligns well with the provision of telephone advice and triage.

Conclusion

This was a relatively large study of US adults 65 years of age or older that utilized the Nurse HealthLine programme. This evaluation was built on standard economic theory. Patients receiving the right care at the right time will likely recover faster and have lower health-care expenditure than those that do not. Those callers following the nurse's recommendation regarding where to receive care were more likely to get the right care at the right time. The impact of nurse adherence was estimated by comparing the average downstream differences in expenditures between compliant and non-compliant callers after removing case-mix differences between the groups via multivariate modelling. While the study was limited to those with an AARP Medicare Supplement Insurance plan provided by UnitedHealthcare insurance company, the authors concluded that the programme was associated with a positive and statistically significant ROI of about $1.59 per dollar spent on the programme. This occurred even though nurses were three times as likely to recommend that callers seek a higher, rather than lower, intensity of care. Future studies should include more details when describing callers who opt not to follow the nurses' recommendations and their subsequent outcomes. Future studies should also attempt to quantify savings and losses for non-triage aspects of the Nurse HealthLine programme.

Acknowledgements

The authors thank Frank G. Bottone, Jr, PhD for his editorial assistance and critical review of this manuscript.

Sources of funding

Funding for this evaluation was provided by United-Health Group.

Ethical approval

This economic evaluation has not been reviewed by an Ethics Board. The evaluation we report in this paper is of a nurse health triage service; such services are offered in most insurance plans. As noted in the manuscript, millions of Americans use these services from many insurance plans across the country. The

NurseLine program for AARP has existed for many years prior to the evaluation we conducted. Members call of their own volition and ask for recommendations regarding where to get care and for educational information to prepare them for that visit. The research team's work was limited to estimating the financial impact of this service using medical claims data. All members of the research team are certified each year by Optum, and each of these individuals has passed a yearly exam that demonstrates knowledge about policies to secure data and handle sensitive information. The research did not involve any additional treatment or services for AARP members. We do not believe this research requires approval by an Ethics Board.

References

Allen-Davis J.T., Beck A., Parker R., Ellis J.L. & Polley D. (2002) Assessment of Vulvovaginal complaints: accuracy of telephone triage and in-office diagnosis. *Obstetrics and Gynecology* **99**, 18–22.

Antos J.R. (2012) The Wyden–Ryan proposal — a foundation for realistic Medicare reform. *New England Journal of Medicine* **366**, 879–881.

Austin P.C. (2011) Comparing paired vs non-paired statistical methods of analyses when making inferences about absolute risk reductions in propensity-score matched samples. *Statistics in Medicine* **30**, 1292–1301.

Barber J.W., King W.D., Monroe K.W. & Nichols M.H. (2000) Evaluation of emergency department referrals by telephone triage. *Pediatrics* **105**, 819–821.

Bogdan G.M., Green J.L., Swanson D., Gabow P. & Dart R.C. (2004) Evaluating patient compliance with nurse advice line recommendations and the impact on health-care costs. *American Journal of Managed Care* **10**, 534–542.

Bunik M., Glazner J.E., Chandramouli V., Emsermann C.B., Hegarty T. & Kempe A. (2007) Pediatric telephone call centers: how do they affect health care use and costs? *Pediatrics* **119**, e305–e313.

Cariello F.P. (2003) Computerized telephone nurse triage. an evaluation of service quality and cost. *Journal of Ambulatory Care Management* **26**, 124–137.

Dent R.L. (2010) The effect of telephone nurse triage on the appropriate use of the emergency department. *Nursing Clinics of North America* **45**, 65–69.

Frisbee S.J., Malloy M., Meurer J.R., Kuhagen K.A. & Kini N.M. (2001) Urban Wisconsin pediatric patients using an after-hours telephone triage service: outcomes and compliance. *World Medical Journal* **100**, 55–58.

Heckman J.J., Hidehiko I. & Prieti T. (1997) Matching as an econometric evaluation estimator: evidence from a job training program. *Review of Economic Studies* **64**, 605–654.

Hildebrandt D.E., Westfall J.M. & Smith P.C. (2003) After-Hours telephone triage affects patient safety. *Journal of Family Practice* **52**, 222–227.

Kempe A., Bunik M. & Ellis J. (2006) How safe is triage by an after-hours telephone call center? *Pediatrics* **118**, 457–463.

Lancet. (2001) Nurse telephone-triage. *Lancet* **357**, 323.

Lee T.J., Baraff L.J., Guzy J., Johnson D. & Woo H. (2003) Does telephone triage delay significant medical treatment? Advice nurse service vs on-call pediatricians. *Archives of Pediatric & Adolescent Medicine* **157**, 635–641.

May D. (2011) The clock is ticking debt talks put bull's-eye on health spending, but how much pain is necessary? *Modern Health-care* **41**, 23.

Melzer S.M. & Poole S.R. (1999) Computerized pediatric telephone triage and advice programs at children's hospitals: operating and financial characteristics. *Archives of Pediatric & Adolescent Medicine* **153**, 858–863.

Moore J.D., Saywell R.M., Thakker N. & Jones T.A. (2002) An analysis of patient compliance with nurse recommendations from an after-hours call center. *American Journal of Managed Care* **8**, 343–351.

Morgan S.L. & Todd J.L. (2008) A diagnostic routine for the detection of consequential heterogeneity of causal effects. *Sociological Methodology* **38**, 231–281.

Nauright L.P., Moneyham L. & Williamson J. (1999) Telephone triage and consultation: an emerging role for nurses. *Nursing Outlook* **47**, 219–226.

O'Connell J.M., Johnson D.A., Stallmeyer J. & Cokingtin D. (2001a) A satisfaction and return-on-investment study of a nurse triage service. *American Journal of Managed Care* **7**, 159–169.

O'Connell J.M., Stanley J.L. & Malakar C.L. (2001b) Satisfaction and patient outcomes of a telephone-based nurse triage service. *Managed Care* **10**, 55–56.

Robinson D.L., Anderson M.M. & Acheson P.M. (1996) Telephone advice: lessons learned and considerations for starting programs. *Journal of Emergency Nursing* **22**, 409–415.

Robinson D.L., Anderson M.M. & Erpenbeck P.M. (1997) Telephone advice: new solutions for old problems. *Nurse Practitioner* **22**, 179–180, 183–6, 189 passim.

Rosenbaum P.R. (1987) Model-Based direct adjustment. *American Statistician* **82**, 387–394.

Rosenbaum P.R. & Rubin D.B. (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**, 41–55.

Rutenberg C.D. (2000) What do we really know about telephone triage? *Journal of Emergency Nursing* **26**, 76–78.

Sabin M. (1998) Telephone triage improves demand management effectiveness. *Health-care Financial Management* **52**, 49–51.

Valanis B., Moscato S., Tanner C. *et al.* (2003) Making it work: organization and processes of telephone nursing advice services. *Journal of Nursing Administration* **33**, 216–223.

Wennberg J.E. & Cooper M.M. (eds) (1999). The quality of medical care in the United States: a report on the Medicare program. The Dartmouth Atlas of Health Care 1999. Published by The Center for the Evaluative Clinical Sciences, Dartmouth Medical School. Available at: http://www.dartmouthatlas.org/downloads/atlases/99Atlas.pdf, accessed 27 April 2012.

Wennberg J.E., Fisher E.S. & Skinner J.S. (2002) Geography and the debate over Medicare reform. *Health Affairs (Millwood)* (Suppl Web Exclusives) W96–W114.

Wolf-Klein G.P. & Silverstone F.A. (1987) A hot-line emergency service for the ambulatory frail elderly. *Gerontologist* **27**, 437–439.