FINAL REPORT OF THE VOLUNTEERS OF AMERICA CONTROLLED PILOT STUDY IN HOMESTEAD AT MAPLEWOOD AND ELDER HOMESTEAD ASSISTED LIVING FACILITIES

1. Methodological Summary of the Conducted Pilot:

In this controlled experimental outcomes study, VOLUNTEERS OF AMERICA and the Medical Automation Research Center (MARC) aimed to evaluate the impact of using unobtrusive remote monitoring technologies to monitor the activities of older adults, including certain health related diagnostic capabilities of the Technology to professional caregivers. The study examined:

1. The economic impact on labor costs associated with
    a. Changes in staff efficiencies in dispatching appropriate services accordingly
    b. Projected financial impact to the caregiver resulting from improved efficiencies and reductions in liability costs and/ or additional revenue stemming from charges for additional care needs identified through the use of the Technology

2. The health outcomes of the residents following monitoring coupled with professional intervention
    a. Determining the utility of the monitored variables to the identification of known and unknown health issues quantitatively
    b. Assessing the diagnostic utility of the system to caregivers qualitatively.

The Abbreviated Specific Aims were:

1- To examine current procedures in the test facility(ies) and jointly develop a diagnostic model linking the variables monitored to disease processes common in elders

2- To assess the impact of the monitoring "Technology" on the productivity of professional caregivers, health related diagnostic utility of the information gathered, and the associated cost at the test facility(ies).

1- Cohorts included monitored Older Adults in two facilities (Homestead at Maplewood, and Elder Homestead), and Professional Caregivers. In addition, an age, gender and overall health status matched group of older adults were selected by the VOLUNTEERS OF AMERICA research study team from a control site (at Coon Rapids) as a control group. Similarly, a control group of professional caregivers was recruited from the control site and enrolled in the study.
2- A focus group was conducted with professional caregivers to solicit information on the diagnostic utility of the parameters reported by the monitoring system with its current sensor modalities. During the focus group a *Diagnostic Information Gathering Instrument* was administered to professional caregivers to collect information that would help to build a model linking changes in monitored parameters to common geriatric diseases.
3- Data on all billable interventions including doctor visits, laboratory tests, hospital days, and Emergency Room (ER) visits, ER days, as well as the frequency of non-billable communications with their health care provider were collected throughout the 3 months pilot period (April 1st- June 30th) for both the monitored subjects and their matching controls for comparison.

4- The *Workload Scale Instrument* was administered to Professional Caregivers in the monitoring and the control sites at the beginning of study as well as the end. In addition, Care Recipient-to-Caregiver ratios were to be assessed at the time of administering the *Workload Scale Instrument*; these are used in normalizing the workload scales to eliminate the effect of possible fluctuations in staffing levels.
5- Finally, a survey, the *Qualitative Assessment of the Utility of the Monitoring System survey*, that asks staff about their opinions in the monitoring technology from Site Accessibility and Ease of Navigation, Site Clarity and Organization, System and Sensor Components Diagnosis, Efficacy, Time Efficacy, and Accuracy perspectives.

2. Study Subjects:

The study involved 21 older adult participants who had the monitoring system installed in their apartments, five males, and 16 females. All participants were over the age of 65 (mean age 88 years +/- 6.4, median age 90, minimum age 73, and maximum age 99). All participants were white. The monitoring group was gender, age, and overall health status matched to a control group of 21 residents from another site in the same metropolitan area that was designated as a control site. All participants were over the age of 65 (mean age 88 years +/- 5.7, median age 88, minimum age 77, and maximum age 97). All participants were white.

Moreover, 12 professional caregivers (nurses and aides) were recruited and enrolled in the study. Six professional caregivers were recruited form the control site, five females and one male, two of whom were nurses and the rest were aides. Mean age of the control site professional caregivers was 41.83 (median age 43, minimum age 33, and maximum age 47). In addition, six professional caregivers were recruited from both monitoring sites (3 from each site), five females and one male, three of whom were nurses and the rest were aides. Mean age of the control site professional caregivers was 46 (median age 47, minimum age 29, and maximum age 63).

3. Results:

3.1. Focus Group:

The focus group and the information collected using the *Diagnostic Information Gathering Instrument* have shown that the bed monitor's information (bed time, time in bed, wake-up time, times up during the night, pulse, breathing, and restlessness), the gait monitor's information (falls, walking velocity, and step length), in addition to the overall number of bathroom visits, were the most significant parameters to monitor. The specific information gathered will allow the development of a **Diagnostic Aid Tool** that will provide caregivers with possible causes of observed changes in the monitored parameters. A preliminary rule-based model was developed. However, the model was not implemented since the validation of the output of this tool requires a longitudinal study, as well as continuous interaction between the development team, clinicians and the professional caregivers that will be using it.

3.2. Frequency of Doctor Visits and other Billable Interventions:

Monitored study subjects (N=21) had a total of 47 doctor visits, including 2 Emergency Room (ER) visits (which does not take into account length of subsequent hospital stay), 16 laboratory tests, and 1 hospital visit (which does not take into account length of stay), with a mean of 2.24 visits per subject (SD=2.256), whereas age, gender, and health matched controls (N=21) had 73 doctor visits, including 11 ER visits, 4 hospital visits, 18 laboratory tests, and 1 urgent care visit, with a mean of 3.48 visits per control subject (SD=2.676). A paired t-test for means has yielded a two-tailed P value of **0.0403**, which is **significant**.

Despite the difference in the overall doctor visits (including frequency of hospital visits, urgent care visits, and visits to and/ or by a physician's assistant) alone, which averaged at 1.38 for monitored subjects (SD=1.322) and 2.10 for control subjects (SD=1.513), between the monitoring and the control groups, an unpaired t-test for means has yielded a two-tailed P value of **0.1112**, which **is not significant**.

Similarly, a comparison using the Mann-Whitney non-parametric test on the number of ER visits alone, 2 for monitored subjects with an average of 0.095 ER visit (SD= 0.301) compared to 11 for the control subjects with an average of 0.52 ER visit (SD=0.750), has yielded a two-tailed P value of **0.0568**, which is **not quite significant**. The non-parametric test was chosen because pairing was not effective and the data points did not have a Gaussian distribution. **Nonetheless, this comparison indicates that the appropriate use of the technology may indeed result in reducing ER visits, which are generally costly, and that statistical significance may be attained in a longer pilot or with a larger sample size.**

The comparison clearly demonstrates the potential cost savings the monitoring technology could bring to payers.

3.3. Hospital Days:

Monitored study subjects (N=21) had a total of 7 hospital days, including 2 emergency room days, with a mean of 0.33 days per subject (SD=1.317), whereas age, gender, and health matched controls (N=21) had 33 hospital days, including 17 emergency room days, with a mean of 1.57 days per control subject (SD=2.226). A Mann-Whitney non-parametric test has yielded a two-tailed P value of **0.0042**, which is considered **highly significant**. This non-parametric test was chosen because pairing was not effective-the data points did not have a Gaussian distribution. Again, the average cost of hospital day stay, in the case of both inpatient and ER stays, is significantly higher than the average cost of a regular visit to the physician.

This comparison provides additional and compelling evidence that the appropriate utilization of the technology could result in cost savings to payers.

3.4. Frequency of Non-Billable Communication with Health Care Provider:

Monitored study subjects (N=21) had a total of 59 calls and faxes, with a mean of 2.81 communications per subject, whereas age, gender, and health matched controls (N=21) had 68

calls and faxes, with a mean of 3.24 communications per control subject. However, **the difference in communication frequency between the groups was not statistically significant**.

This shows that the technology did not reduce billable intervention through the reliance on non-billable interactions with the physician/ health care provider, but rather through possible preventive measures and proactive interventions provided by the facility caregivers. Hence, the technology, when utilized appropriately, should not result in added communications burdens on the external health care provider. This is an important factor in the overall acceptance of the proliferation of this monitoring technology by physicians and other health care professionals.

3.5. Comparison on Overall Cost of Care to Payer:

This comparison was based on pricing and average cost information obtained from U Care Minnesota, a leading independent nonprofit Health Maintenance plans provider in the Minneapolis metropolitan area. The average cost of physician visits, urgent care visit, and lab test was estimated at $65.75 per visit; and average transportation cost for round trip in a taxi is estimated at $30.0 private insurance pay per visit (other transportation options available are $3 round trip reimbursable by Medicare/ Medicaid, and $6 average cost for round trip by a family member). The average cost of an ER visit that does not entail overnight stay was estimated at $296.64 per visit, whereas the average cost of an overnight stay following an ER visit, as well as regular hospital day stay, was estimated at $1884.86 per day. Using the above information/ assumptions, and the frequency of doctor visits, laboratory tests, ER visits and hospital day stays, the **total cost of care for the monitoring group was calculated at $17,407.02 with an average cost of $828.91** (SD=2458.90) **per monitored individual**, **compared to $67,753.88 f**or **the control group, with an average cost of $3236.38** (SD=4214.0) **per control individual**, **without taking the cost of medications into account**, since medications data was not collected. The difference in the cost of care to payer between the two cohorts was **statistically very significant**, as shown by the Mann-Whitney non-parametric test result which has yielded a two-tailed P value of **0.0046, despite the relatively small sample size (N=21 in each group), and the short period of the pilot (only 3 months)**. The comparison did not take the cost of monitoring into account either. The non-parametric test was chosen because pairing was not effective and the data points did not have a Gaussian distribution.

Assuming a monitoring fee of $60.00 per month per monitored individual, the cost for the monitoring group becomes $21,187.02, and the average cost per monitored individual becomes $1008.91, compared to $67,753.88 for **the control group, with an average cost of $3236.38**. The difference in the cost of care to payer between the two cohorts remained **statistically significant**, as shown by the Mann-Whitney non-parametric test result which has yielded a two-tailed P value of **0.0346**. The non-parametric test was chosen because pairing was not effective and the data points did not have a Gaussian distribution. **Moreover, assuming each monitoring system costs a $1000.00 (which is the approximate cost of the current research prototype) and is covered by the payer, the investment in the system's initial cost is recaptured in less than six weeks of deployment of the monitoring system in care cost savings.**

This comparison provides additional and compelling evidence that the appropriate utilization of the technology could result in cost savings to payers.

3.6. Professional Caregiver Workload Scales:

At the beginning of the study Professional Caregivers at the monitoring sites (three caregivers from Homestead at Maplewood and 3 caregivers from Elder Homestead, N=6) had an average score of **8.83** on the workload scales (SD= 2.401), where as caregivers at the control site (N=6) had an average score of **10.17** (SD= 4.875). The difference in workload scales was not significant, as shown by the Mann-Whitney non-parametric test result, which has yielded a P value of 0.6882 (**i.e. not statistically significant**). However, the two monitoring sites had Care Recipient-to-Caregiver ratios of **14.11** and **11.78** respectively, compared to only **6.62** for the control site; there are two separate buildings in the control site, one dedicated for memory care residents, and this requires higher staffing levels, and hence lower care recipient to caregiver ratios. **It is worthy to note that there are two separate buildings at the control site, with one dedicated for memory care residents, and this may be the reason for higher staffing levels, and hence lower Care Recipient to Caregiver ratios.** Since the Care Recipient-to-Caregiver ratio represents in essence caregiver efficiency, it may have a significant effect on the workloads of caregivers, and since this ratio may change throughout the period of the study, we measured these ratios at the times we collected workload scale data and used these ratios to normalize the workload scale scores. The efficiency normalized workload scale scores of Professional Caregivers at the sites that utilized the monitoring technology had a mean of **0.6** (SD= 0.208), compared to a mean normalized scores of **1.38** (SD= 0.662) at the control site. A Mann-Whitney non-parametric test has yielded a two-tailed P value of **0.0411**, which is considered **significant**.

After three months, and at the end of the pilot, the workload scale instrument was re-administered to caregivers, and the Care Recipient-to-Caregiver ratios were collected from the respective sites at the time of re-administration of the workload scales. Professional Caregivers at the monitoring site scored a mean of **10.33** (SD= 3.44) on the workload scale instrument, compared to a mean of **12.83** (SD= 3.25) for the control site. Care Recipient-to-Caregiver at one of the monitoring site has dropped slightly from 11.78 to **10.89**, and remained at **14.11** at the other. This ratio remained the same for the control site at **6.62**. The normalized workload scale scores of Professional Caregivers at the monitoring site had a mean of **0.84** (SD= 0.282), compared to a **1.94** (SD= 0.491) at the control site. A Mann-Whitney non-parametric test has yielded a two-tailed P value of **0.0022**, which is considered **very significant**. The slight increase in workload scale scores, as well as normalized workload scale scores, over the course of the study at both the monitoring and the control sites may be due to additional burdens the research study may have exerted on participating Professional Caregivers including the focus groups, the administered research instruments, etc. However, **these changes were not statistically significant.** The non-parametric tests were chosen here due to the small sample size.

The final scores of workload scale and the normalized workload scale at both the monitoring and the control sites show that professional caregiver workloads at the sites that utilized the monitoring technology were significantly lower than the workloads of peers at the control sites who were <u>not</u> using the technology, <u>despite the fact that the former were serving significantly larger number of care recipients</u>. The normalized

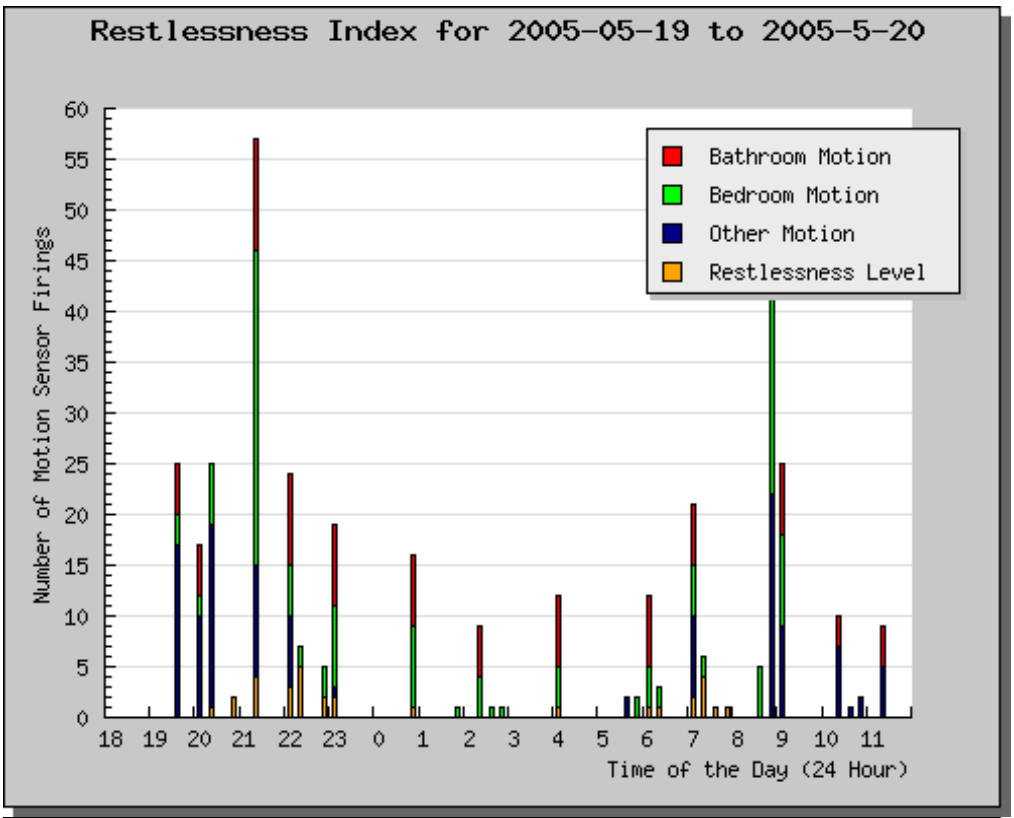
workload scale, in essence, takes into account staff efficiencies. Thus the study showed that the caregivers at the monitoring sites have attained significantly higher levels of efficiency than their control site counterparts, which could be due to the effective utilization of the monitoring technology over the past year. The difference in efficiency normalized workload scales between the sites are statistically significant despite the very low sample size (N=6 at each site).

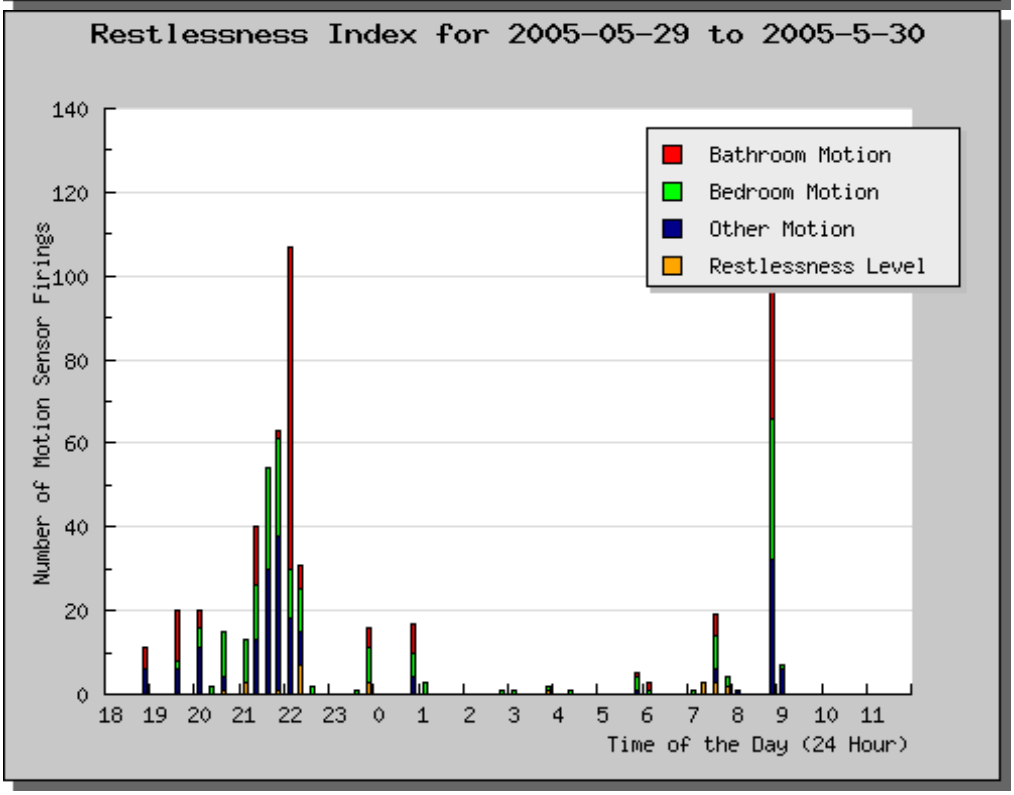
4. Alerts:

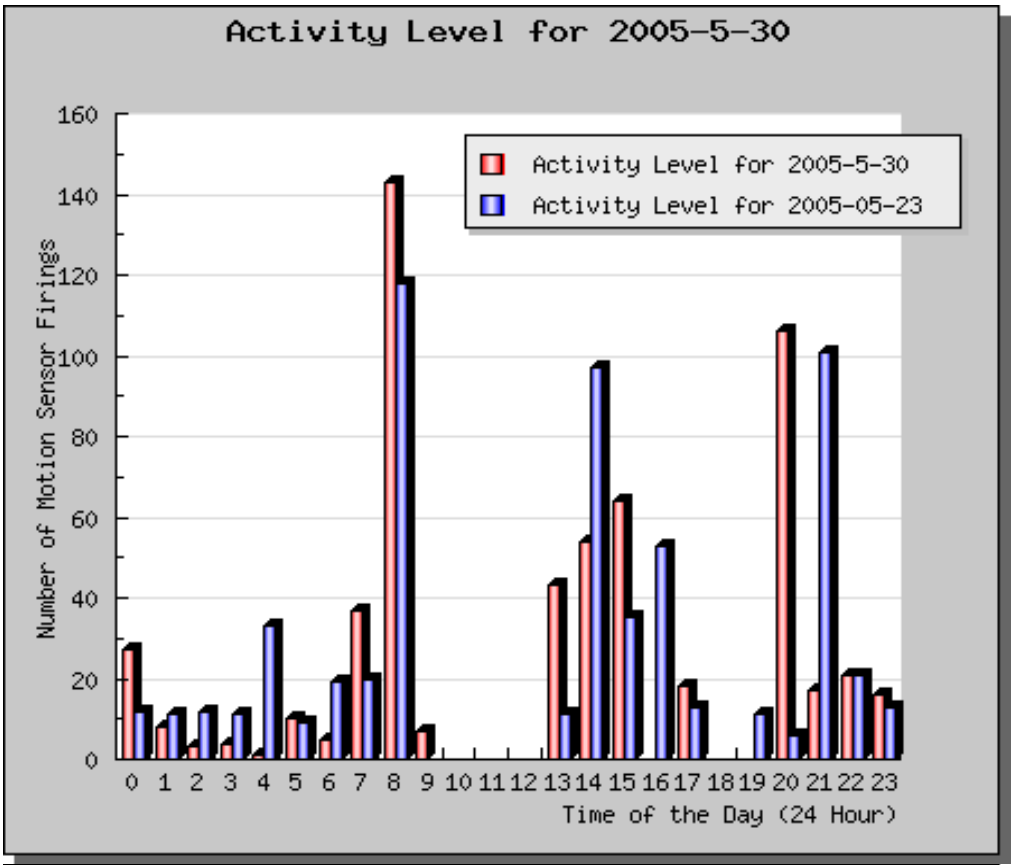
The alerts missed two actual falls, one in each facility, and the residents were discovered by staff before the pager went off. Also the alerting sub-system generated a significant number of false alerts, especially in the beginning of the study. This was due to changes made to the alert routines to accommodate the pilot that was running in parallel. The majority of the pilot period was considered a refinement for the alerts. Towards the end of the pilot, facility managers at the monitoring sites expressed their satisfaction with the significant reduction in false alerts attained.

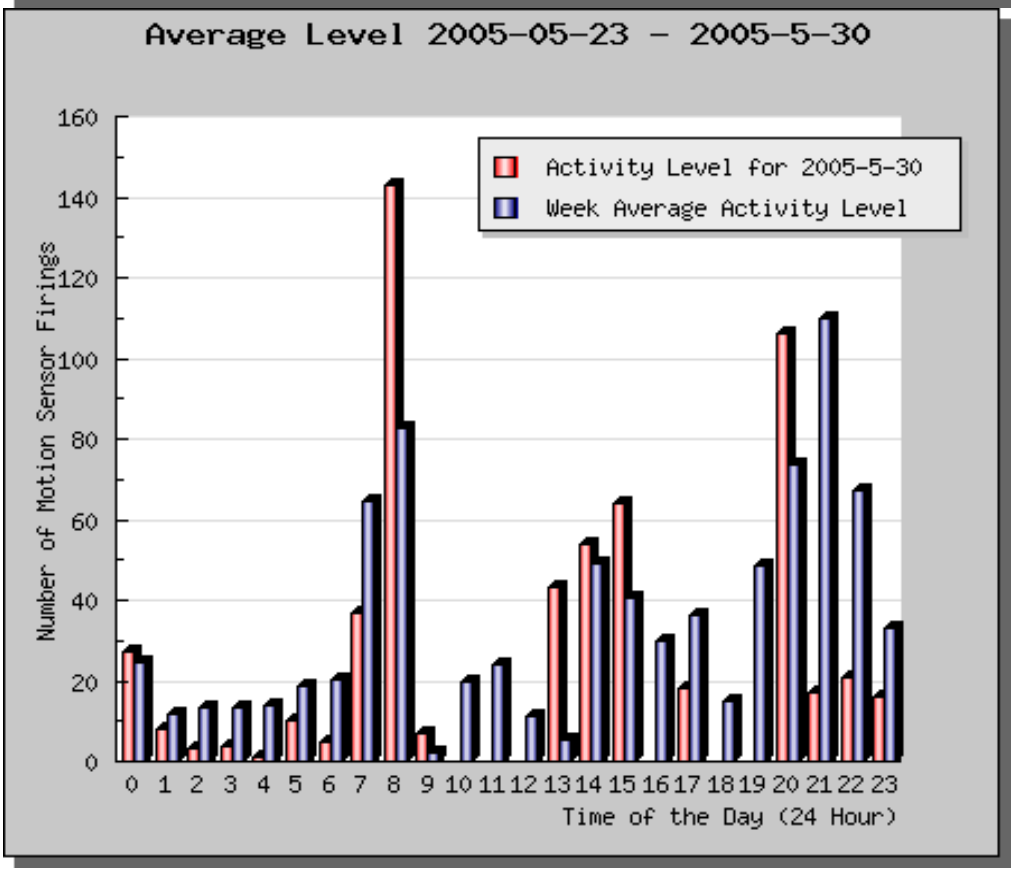
5. Case Study:

Subject 1127 had excessive bathroom visits during the night of the May 20th (above her nightly average), and was up several times during the night. Facility management and professional caregivers suspected a Urinary Tract Infection (UTI), and advised to see the physician who ordered a Urine Analysis (UA); however, UA results were negative. After further investigation prompted by the abnormalities in her activity patterns on her monitoring reports, the Caregivers at the Elder Homestead facility made additional inquires of the resident. In response to this caregiver probing, the subject volunteered she was feeling pain, whereupon the staff decided to put the resident on Tylenol 3. As a result of the intervention, the number of nightly bathroom visits was reduced from 6 to 3, and restlessness and nightly activity levels were considerably reduced as the graphs below show.


Restlessness Index for 2005-05-19 to 2005-5-20
Number of Motion Sensor Firings
Time of the Day (24 Hour)
Bathroom Motion
Bedroom Motion
Other Motion
Restlessness Level
0
5
10
15
20
25
30
35
40
45
50
55
60
18 19 20 21 22 23 0 1 2 3 4 5 6 7 8 9 10 11


Restlessness Index for 2005-05-29 to 2005-5-30
Number of Motion Sensor Firings
Time of the Day (24 Hour)
Bathroom Motion
Bedroom Motion
Other Motion
Restlessness Level
0
20
40
60
80
100
120
140
18 19 20 21 22 23 0 1 2 3 4 5 6 7 8 9 10 11


Activity Level for 2005-5-30
Activity Level for 2005-5-30
Activity Level for 2005-05-23
Number of Motion Sensor Firings
0
20
40
60
80
100
120
140
160
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
Time of the Day (24 Hour)


Average Level 2005-05-23 - 2005-5-30
Activity Level for 2005-5-30
Week Average Activity Level
Number of Motion Sensor Firings
0
20
40
60
80
100
120
140
160
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
Time of the Day (24 Hour)

6. Conclusions:

1- *a.* The focus group discussions amongst professional caregivers at the monitoring as well as the control site revealed that bed monitor's information (bed time, time in bed, wake-up time, times up during the night, pulse, breathing, and restlessness), and the gait monitor's information (falls, walking velocity, and step length) in addition to the number of bathroom visits were the most significant parameters to monitor. *b.* Based on the information reported by these monitoring devices/ components it is possible to infer possible causes and present those as candidate diagnoses to aid the professional caregiver. *c.* the implementation of the Diagnostic Aid Tool is possible, but requires a longitudinal testing and validation pilot with significant input from clinical staff as well as clinicians at the University.
2- The technology could result in significant cost savings to payers, as shown by the statistically significant difference in doctor visits, ER visits, hospital days and other billable interventions between the monitored group and the control group over the course of the pilot, despite the short study period.
3- There was no evidence that the technology might have caused physicians of monitored patients additional non-billable communications burdens.
4- The workload scale and Care Recipient-to-Caregiver ratios indicated that the monitoring technology may result in lower workloads and significantly higher caregiver efficiencies simultaneously.
5- The case study demonstrated, once more, the utility of the technology as a proactive health tool that positively affects the quality of life of the monitored individuals.