Published in final edited form as:

J Clin Epidemiol. 2011 March ; 64(3): 320–330. doi:10.1016/j.jclinepi.2010.02.019.

Weighted index explained more variance in physical function than an additively scored functional comorbidity scale

Linda Resnik, PT, PhD, OCS, **Pedro Gozalo, PhD**, and **Dennis L. Hart, PT, PhD**

Abstract

Objective—1) examine association between the Functional Comorbidity Index (FCI) and discharge functional status (FS); 2) examine impact of FCI on FS when added to comprehensive models, and 3) compare additive FCI to Weighted FCI and list of condition variables (list).

Study Design and Setting—Patients drawn from Focus On Therapeutic Outcomes, Inc. (FOTO) database (1/1/06-12/31/07). FS collected using computer adaptive tests. Linear regression examined association between FCI and FS. Three methods of including functional comorbidities (FC) were compared.

Results—Relationship between FCI and FS varied by group (range .02 0.9). Models with weighted index or list had similar R^2. Weighted FCI or list increased R^2 of crude models by < .01 for cervical, shoulder, and lumbar; .01 wrist/hand, knee and foot/ankle; by.02 hip; 03 elbow; and .08 neurological. Addition of FCI to comprehensive models added <.01 to R^2 (all groups). Weighted FCI increased R^2 by <.01 for cervical, lumbar and shoulder,.01 wrist/hand, hip, knee, foot/ankle, .02 elbow, and .04 neurological; whereas list increased R^2 by < .01 cervical, shoulder and lumbar, .01 knee and foot/ankle, .02 elbow, wrist/hand and hip and .05 neurological.

Conclusion—List of comorbidities or weighted FCI is preferable to using additive FCI.

Keywords

Risk adjustment; comorbidity; Outcome Assessment (Health Care); rehabilitation; physical therapy; research design; computerized adaptive testing

---

Outpatient therapies are an important and costly health service, with estimated costs rising at 15 percent per year.[1] In 2000, 3.6 million (8.6% of all) Medicare beneficiaries received outpatient therapy, with total costs amounting to over 2 billion dollars. [1] In 2002, 22 million (2.5 percent of all) visits to physicians resulted in a referral to physical therapy. [2] The cost of these services has prompted several cost-control initiatives and demonstration projects. [3] [4] [Linda, check your EndNote to make reference citation [3–4].]Promising alternatives include development of provider quality profiles and value-based purchasing. [5]

Methods to determine effectiveness (i.e., patient outcomes) and efficiency (i.e., use of services) of out-patient rehabilitation services are needed. Improvement in functional status

---

Corresponding Author: Linda Resnik, PhD, PT, OCS, Research Health Scientist, Providence VA Medical Center, Associate Professor, Department of Community Health, Box G-S121(6), Brown University, Linda_Resnik@Brown.edu, Phone: 401-86309214, Fax: 401-863-3489.

(FS) is a common primary outcome of outpatient therapy. [6–8] Reliable, valid and responsive measures of FS exist, [9–11] but any outcomes measure must be statistically risk-adjusted to control for potential confounders for meaningful interpretation. Hence, the success of any cost methods based upon patient FS outcome depends on the development of precise risk-adjustment methods that adequately estimate the variation in FS and service utilization related to patient factors, so the remaining differences in outcomes can be attributed more directly to the care delivered. [12] Otherwise, providers treating sicker patients may be penalized when patients fail to show enough improvement or use more visits in a treatment episode, [13] which may lead to care access barriers.

Risk adjustment is critical for evaluating outpatient health outcomes because differences in patient FS are associated with type and severity of impairments as well as many other factors like comorbid conditions, [14–16] all of which can affect FS in patients undergoing rehabilitation. Clearly, these variables must be controlled before a meaningful interpretation of FS change can be made. [17]

Currently, there is no consensus on the best risk-adjustment methods for outpatient rehabilitation and little work has been done in this area. However, risk adjustment for case-mix, (i.e., complexity and diversity of patients) is particularly important in observational studies, where treatment groups are self-selected rather than randomized. Previous comparisons between health care providers have been compromised by inadequate assessment of case-mix. [18,19]

Existing indices, such as the Charlson Comorbidity Index [20] and the Kaplan-Feinstein Index, [21] have not been available in outpatient rehabilitation datasets. More importantly, they were developed to predict mortality and as such, they exclude diagnoses such as arthritis, which impact FS but are unlikely to result in mortality. Therefore, it is not surprising that a prior evaluation of indices designed to predict mortality concluded that such comorbid indices had little relationship with physical disability. [22] Thus, researchers investigating FS as the outcome of interest in outpatient rehabilitation generally do not include measures of comorbidity in their risk-adjustment models.

The recently developed Functional Comorbidity Index (FCI), [22] and its addition to the largest available dataset of outpatient rehabilitation, Focus On Therapeutic Outcomes, Inc. (FOTO) offers the potential to improve risk-adjustment methods in this population. The FCI is the only known index designed specifically to control for comorbid conditions that are hypothesized to affect functional status rather than mortality. The index was developed to explain physical function as measured by the SF-36 using two databases; the Canadian Multi Center Osteoporosis study (CaMos) and the National Spine Network (NSN). Both databases included patient self-reported presence of comorbid conditions. [22] Groll et al found the use of a weighted index added little to the predictive ability of the index (adjusted R^2 .30 vs. .28), thus they recommended that the FCI be additively scored. Thus, the FCI is scored by adding the number of “yes” answers to indicate the history of specific conditions. A score of 0 indicates absence of any comorbid conditions, and a score of 18 indicates the highest number of comorbid conditions contained in the FCI. Groll reported that the FCI showed a stronger association with FS, as measured by the SF-36 PF-10 scale, as compared with the Charlson or Kaplan-Feinstein indices, which were developed to predict mortality. [22]

A critical consideration for additive indices is whether the effect of specific comorbid conditions on FS varies substantially by type of body impairment for which they seek rehabilitative treatment, in which case a summative scoring method may not be appropriate for all groups. For example, the effect of obesity (BMI >30) may be negligible for patients with wrist/hand impairments, but more substantial for patients with impairments of the lower extremity or lumbar spine. In these instances, use of weighted scoring specific to the impairment category might be preferable. The FCI was developed in cross-sectional studies using the Canadian Multicentre Osteoporosis Study (CaMos) [23] and the National Spine Network (NSN) databases. [24] The National Spine Network database contains data on patients who consulted spine centers due to a variety of spinal problems. Because low back problems are the most prevalent type of spinal problem, we would expect the performance of the additively scored FCI would be best in the population of patients seeking care for low back impairments. Prior research [9,10,11] suggests additive scoring of the FCI discriminates patients with lumbar spine or lower extremity impairments by different levels of functional status. However, no prior research has examined the use of FCI in more diverse outpatient rehabilitation populations, or examined alternative methods of scoring the FCI such as creation of a sub-group specific weighted index.

We are aware of only one study that has examined the relationship between FCI and discharge FS in a multivariate longitudinal analysis. [4] In their preliminary analyses, it appeared the FCI added a small but statistically significant percent of variance controlled in the model R^2 (0.018 to the model). In addition, in univariate analyses, Hart et al demonstrated in patients with lumbar spine or lower extremity impairments that as the number of comorbid conditions increased, discharge FS decreased. [9,11,25] Because the FCI was developed to be a predictor of FS using cross-sectional data and preliminary data to suggest more comorbid conditions will decrease discharge FS, we wanted to test whether similar relationships will persist in the context of more robust longitudinal analyses. It is possible that the relationship between FCI and intake FS may be confounded when patients have been referred to therapy services because of the consequences of the onset or exacerbation of one of the chronic conditions measured by the FCI. We believe that examination of the relationship between FCI and discharge FS may more accurately capture functional status and would be more relevant for case-mix adjustment. Furthermore, we do not know if use of the FCI would improve predictive models of discharge FS when added to comprehensive models that we have used in our prior research which have contained variables such as intake FS, age, onset of condition, surgical history, gender, and type of insurer. [8,26–28]

In summary, the FCI is a promising new comorbidity index for use in outpatient rehabilitation that might be important when predicting FS outcomes. However, further validity and utility testing is needed before its use in risk-adjustment models for a wide variety of outpatient rehabilitation patients with varying conditions can be recommended. Thus, the purposes of this study were to 1) examine the association between the Functional Comorbidity Index (FCI) and functional status (FS) at discharge among patients seeking outpatient rehabilitation using a variety of methods for calculating FCI score, and 2) examine the impact of FCI on discharge FS when added to previously used comprehensive predictive models of FS.

Methods

This was a prospective, longitudinal, cohort study. Patients were followed from intake to rehabilitation services through discharge. The study population was drawn from the Focus On Therapeutic Outcomes, Inc. (FOTO) database of patients undergoing outpatient rehabilitation. [29,30] FOTO included the Functional Comorbidity Index as a mandatory module for all patients in May 2005. Our cohort consisted of all individuals (N=101,418) entered into the FOTO database with an intake date between January 1, 2006 and December 31, 2007, who had completed FS assessment by computerized adaptive testing (CAT) methods (see below) for one of the nine impairment groups. Of these, we eliminated those who had missing discharge FS data (N=38,995), those who had missing data on one or more

FCI variables (N=10,906), those who had missing any other variable used in our analyses (N=6,592), leaving a final analytic sample of 44,925.

The FOTO database, the largest outpatient rehabilitation database available for researchers in the United States, provides a standardized set of data collection instruments that contains demographic, intake and discharge functional status (FS) measures (described below), and administrative data from outpatient rehabilitation services. Whereas regulatory requirements mandate the collection of clinical outcomes data in skilled nursing facilities through the minimum dataset (MDS), [31] home care agencies through the use of the Outcome and Assessment Information Set (OASIS), and inpatient rehabilitation through the use of Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI), there are presently no similar requirements for outpatient rehabilitation settings, thus limiting the availability of population-based datasets. While no national database of outpatient rehabilitation outcomes comparable in size to the MDS, OASIS or IRF-PAIs exists, many outpatient facilities track patient data in their own data collection system, or with the assistance of an external outcomes measurement provider, like FOTO. [32,33]

Mode of data collection

When a clinic starts collecting FOTO data, staff provide information describing the clinic, and clinic staff are trained in the data collection process. All patients completed self-report functional status surveys via CAT administrations prior to their initial evaluation and following discharge from their rehabilitation episode. Patient demographic and FCI data were collected at intake.

Key Independent Variable: Functional Comorbidity Index

The Functional Comorbidity Index (FCI) is the only known index designed with physical function as the explicit outcome of interest. The FCI uses patient self-report of the comorbid conditions, which correlates well with data collected by medical record review. [22,34] The FCI contains the following eighteen condition variables: arthritis, osteoporosis, asthma, chronic obstructive pulmonary disease (COPD), angina, congestive heart failure (CHF), prior heart attack (MI), neurological disease (multiple sclerosis, Parkinson’s disease), prior stroke or transient ischemic attack (TIA), peripheral vascular disease (PVD), diabetes, upper gastrointestinal, disease (including ulcers and gastroesophageal reflux), depressed mood, anxiety, visual impairment, hearing impairment, obesity (calculated from BMI >30), and low back pain; a score of 0 indicating absence of any comorbid illness, and a score of 18 indicating the highest number of comorbid conditions. In the FOTO system, patients identify from a list of medical problems containing 17 FCI conditions (absent body mass index (BMI)) any problem that applies to them. Patients provide their height and weight from which their BMI is calculated. Presence of obesity (the 18th FCI) is identified if BMI is 30 or above. No other measure of comorbidity (i.e., medical record abstraction or interview) was used to collect FCI data.

Dependent Variable

For the purpose of this study, the outcome of interest is discharge functional status (FS), which we define as the patient’s perception of their ability to perform functional tasks described in the FS items. FS is of interest because many people seek rehabilitation to improve functional deficits, and patient self-report of FS has become a well accepted outcome in research and clinical practice, particularly in outpatient rehabilitation. [7,15,16,35] FS cannot be directly observed but can be estimated by analyzing patient’s responses to a set of items. [36,37] FS data are collected using 9 CAT administrations: one for each of the body impairment categories under study (Table 1). Patients with assistance from staff identified which body part (or neurological impairment) was their primary reason for rehabilitation, which directed the computer to administer one CAT per patient. Each CAT generated FS measures specific to the patient's impairment. Each CAT administration produced an estimate of FS that was transformed to a 0 (low functioning) to 100 (high functioning) metric, but it should be noted that the mathematical equivalence of the FS estimates across the body part impairments has not been tested necessitating separate analyses per body part specific FS measure. For each CAT administration, the first item administered is the median level difficulty item for the specific impairment. Estimates of the patient's FS ability with their standard error (SE) are estimated using a maximum likelihood estimation routine after each item is answered. [38] Subsequent items are selected by the computer by matching the current estimate of the patient's FS to a previously un-asked item with the most information at the current estimated patient's FS. The CAT stops if either the SE for the current FS estimate is less than 4 out of 100 FS units or the change in the last three FS estimates are each less than one out of 100 FS units.

Hart et al have described the development of each CAT including the discriminant validity of the CAT estimated FS measures and the operating characteristics of each CAT for lower limb impairments, [9–11,39] shoulder impairments, [40] and lumbar impairments. [25,41] Each CAT used for elbow, wrist/hand, cervical regions, or neurological conditions, which have been described in a report on pay-for-performance, [4] contains items from the SF-12, [42] SF-36, [43] and items pertinent to patients with upper extremity impairments, items representing lower functional abilities and items pertinent to specific impairments. [37] The resultant 50 items were calibrated into an FS scale using a rating scale IRT model. [44] Then, clinicians placed the items into groups related to the patient's impairment, including body part treated or neurological condition. For example, if the patient had a cervical impairment, only items pertinent to the cervical area were used in the condition-specific CAT, or if the patient had a neurological condition (peripheral nerve impairments, injuries to nerves, traumatic brain injuries, cerebral vascular accidents, and other), only items pertinent to a neurological condition were used in the condition-specific CAT. The FS measures estimated from each CAT have data supporting reliability, validity, and responsiveness. [4] The responsiveness, sensitivity to change, construct validity and clinical interpretation of the FS measures generated by these condition-specific CAT administrations have been reported as strong [4,9,10,25,40,45,46] as was the usability and efficiency of the CAT administrations. [46]

Although CAT administrations have been in existence for over 25 years, [47] the CAT method of FS data collection is new to rehabilitation. [9,11,40,41] CAT offers advantages compared to a computer administered or paper and pencil instruments. CAT match item difficulty to the patient's level of ability reducing the number of inappropriate items administered; administer fewer items reducing respondent burden; and allow the level of measure precision to be established before testing. Therefore, CAT facilitates good measurement precision with low response burden. [9,40,41,49]

Preliminary analyses of biased censoring

In order to use FOTO data to model outcomes of FS and number of therapy visits, data must be collected at several points in time (i.e., intake and discharge from therapy). If patients discontinue therapy and no FS data are collected upon discharge those patients are considered censored. In these analyses, we considered patients censored if they had missing FS discharge data, FCI data, or any of the other covariates used in our analyses. We compared baseline characteristics (from the intake assessment) for those who were and those who were not censored. Overall, we found small differences. The patients in our sample (those who were uncensored) had slightly higher discharge FS (65.4 vs. 63.8) and were marginally older (52.1 vs. 51.7 years) than those who were censored. Other differences between groups were even smaller in magnitude. To account for potential informative censoring bias, we used inverse probability weighting regression methods to adjust for these observed differences. [50] Inverse probability weighting was accomplished by performing a two-step procedure. In step one, we fitted a logistic regression model where the dependent variable took the value of 1 if the observation was complete, and 0 if missing, and where all patient baseline variables were included as covariates. In step two, we used the inverse of the predicted probabilities of being complete as weights for the patient data. [51] Thus, patients that, based on their data, were unlikely to have complete data were given more weight in estimating FS compared to patients with higher probability of having complete data.

Data Analysis

We analyzed 9 FS measures, each generated using a body part- or condition-specific CAT. Because the mathematical equivalency of the CAT FS measures across impairments and conditions has not been studied, and patients with different types of problems respond differently to some FS items, [39] we analyzed the subgroups of patients separately. Thus, we examined the FCI's association with discharge FS for each patient group.

We examined the association between FCI and discharge FS estimating the FCI score using three methods. First, we modeled discharge FS using the list of 18 condition variables as independent variables. Second, we modeled discharge FS using the additive score of the FCI as the independent variable.

Last, we developed a weighted FCI index, which was sub-group specific and used the score of this index as the independent variable to model discharge FS. We followed the approach of Lubotsky and Wittenberg in developing the weighted index. [52] In this approach, each comorbidity is interpreted as an indirect measure of the common underlying health status. The method allows us to extract and combine the information on health status contained within each separate comorbidity to obtain an improved estimate of the unobserved health status. Note that a separate index needs to be constructed for each impairment group since different comorbidities may have different relative information on the common underlying health status influencing discharge FS for that impairment. To create the index, we estimated linear models of discharge FS for each impairment group using the full list of 18 comorbid conditions as independent variables. These impairment-specific coefficients were then weighted (with the weights being proportional to the inverse of the sum of each coefficient estimate times the covariance between the comorbidity of that coefficient and the outcome) to obtain an impairment-specific index that provides the least unbiased estimate of the underlying health status for that impairment group.

Thus, separate models (27 in total) examining the impact of FCI (as additive FCI, weighted FCI, or long list) on discharge FS were developed for each of the 9 subgroups of patients. We used linear regression (with inverse probability weighting for censoring) in Stata 9. [53] The long list of condition variables can be considered the best choice to maximize predictive ability of the model as use of the list of separate conditions does not force the relevance of each condition to be identical to each other. We compared the R^2 generated from these models as a way to estimate the explanatory power of each of the scoring systems.

Next, we examined the impact of FCI on discharge functional status when added to other predictive models of functional status, testing the hypotheses that a) the addition of the FCI (in any form) to a model containing intake FS would explain significantly more variance, and b) models including the FCI and other case-mix variables would explain more variance than models without the FCI. Separate models examining the impact of FCI on discharge FS were developed for each subgroup of patients. We compared models containing intake FS with models containing intake FS and each of the three scoring methods of the FCI. We also compared comprehensive models (controlling for intake FS, age, gender, onset of condition, exercise history, payer type and surgical history) with comprehensive models (with same independent variables) plus each of the three scoring methods of the FCI.

We examined the relationship between the additively scored FCI and sub-group specific weighted FCI scores by creating scatter plots. We checked the validity of our subgroup specific weighted indices by selecting a random sample of 70% of cases of patients for estimating the weighted index, and then tested the index on the remaining 30% as a validation sample. We performed one hundred such replications, and then compared the R^2 from the weighted index to the other models using the FCI and the list of condition variables.

Results

The FS, FCI scores and other characteristics of the cohort used in this analysis are shown in Table 1. Briefly, patients in the cervical and hip groups had the highest (67%) and the neurological group had the lowest (53%) percentage of women. Patients in the neurological impairment group had the highest number of comorbidities (3.14) and those in the elbow impairment group had the lowest (1.59). Table 2 shows that the prevalence of specific comorbidities varies substantially by impairment group. For example, 49% of patients in the knee and 48% of patients in the hip group had the diagnosis of arthritis as compared to only 24% of patients in the elbow group and 28% of patients in the foot/ankle group. Eight percent of patients in the neurological group had a diagnosis of MI as compared to 2% in the cervical, elbow and 3% in the knee groups. About twice as many people with neurological conditions had diabetes compared to people in the other 8 groups.

Relationship of FCI and weighted FCI

We examined the relationship between the additively scored FCI and sub-group specific weighted FCI scores by creating scatter plots. Two comparisons are shown to illustrate the findings (Figure 1). These plots show that for each value of the additively scored FCI, there may be substantially different values for the weighted FCI. The range of Pearson Correlations is lowest for the neurological (0.55) and elbow groups (0.65) to highest in the lumbar (0.91) and shoulder groups (0.93).

R^2 of models

Results of crude models using only FCI in any scoring format to predict discharge FS showed that the relationship between FCI and FS was small and varied by patient group (Table 3). Overall, the strongest relationships were observed in patients in the lumbar, hip, knee and neurological groups. The R^2 of models using the additively scored FCI to predict discharge functional status varied from a low of 0.02 (elbow group) to high of 0.09 in the hip group. The R^2 of models using either the weighted index or the separate condition variables varied from a low of 0.03 in the shoulder group to a high of 0.14 in the neurological group.

Overall the R^2 of the models using additive FCI scoring were lower than the R^2 of models using weighted FCI or separate condition variables (Table 3). The R^2 of models using the weighted FCI and those using the separate list of condition variables were virtually identical. The increase of explanatory power affected some groups more than others. For example, use of the weighted FCI or list of separate comorbid conditions as compared to use of the additively scored FCI increased the R^2 by < 0.01 points for cervical, shoulder, and lumbar groups; by 0.01 points for the wrist/hand, knee and foot/ankle groups; by 0.02 for the hip group; 0.03 for the elbow group; and 0.08 for the neurological group.

Contribution of FCI to models containing Intake FS

The FCI (either additively scored or as a weighted index) was statistically significant in models containing intake FS for all groups (P=<0.01). Table 4 shows that the additively scored FCI added 0.01 for wrist/hand, neurological, cervical and elbow groups, and approximately 0.02 for shoulder, lumbar, hip, foot/ankle and knee groups to models containing intake FS only. In contrast, addition of the weighted FCI to models containing intake FS only contributed 0.01 for cervical, and shoulder groups, and 0.02 for elbow, wrist/hand and lumbar groups, 0.03 for hip, knee, and foot/ankle groups, and 0.05 for the neurological group. Results for models using the 18 separate conditions variables were similar to models using the weighted index, except for patients in the neurological subgroup where the explanatory power of the model using the list of variables was increased by 0.07 points.

Contribution of FCI to comprehensive models predicting discharge FS

Comprehensive models controlling for intake FS, age, gender, onset of condition, exercise history, payer type and surgical history, but without a measure of FCI, had higher R^2 than simpler models containing both intake FS and any single measure of FCI (Table 4). The addition of the additively scored FCI to comprehensive models controlling for intake FS, age, gender, symptom onset, exercise history, payer type and surgical history did not increase the R^2 for any group by more than 0.01. In contrast, addition of the weighted FCI to comprehensive models increased the R^2 by <0.01 for the cervical, lumbar and shoulder group, 0.01 for wrist/hand, hip, knee, foot and ankle groups, 0.02 for the elbow group, and 0.04 for the neurological group. Addition of the separate list of condition variables increased the R^2 of the comprehensive models by less than 0.01 for the cervical, shoulder and lumbar groups, 0.01 for the knee and foot/ankle groups, 0.02 for elbow, wrist/hand and hip groups and 0.05 for the neurological group.

Validation of the sub-group specific weighted indices

We conducted analyses which confirmed the results obtained from models using the full sample. In all cases, models containing the full list of comorbid conditions had the highest R^2. For 8 out of 9 groups, models containing the weighted FCI had greater explanatory power than did models containing the additively scored FCI. For the remaining group (shoulder impairment), sensitivity analyses suggested the weighted index had R^2 marginally lower than the additively scored FCI.

Discussion

Clinicians argue the need to use measures of comorbid conditions to risk-adjust patient outcomes because they believe that it is more difficult for patients to achieve improvements in functional status when they have more comorbid conditions, which appears logical. The primary purpose of our study was to assess the utility of adding measures of functional comorbidity to models for discharge functional status in a sample of patients receiving outpatient rehabilitation. We explored the use of three separate methods for estimating the effect of having more functional comorbid conditions; the additive FCI score as proposed by Groll, a sub-group specific weighted index, and use of separate condition variables contained in Groll's FCI.

R^2 of models

Our analyses suggested that adding functional comorbid conditions using any scoring method helps to explain discharge functional status. The addition of functional comorbid conditions contributed between 0.2 and 0.14 percent of the variance in crude models. Our findings about the additive FCI are consistent with previous univariate analyses using the FOTO database that reported that discharge FS decreased as the number of comorbid conditions increased for patients with hip, [10], knee, [11] foot or ankle [9] and lumbar spine impairments. . [9,11,25] However, the explanatory value of adding functional comorbid conditions varied by patient impairment group and the scoring methodology we used. For patients in the cervical, shoulder, lumbar, wrist/hand, knee and foot/ankle groups, the explanatory power of models was virtually equivalent for the three methods of scoring functional comorbidities. However, there were differences in explanatory power for models for the hip, elbow and neurological groups. Models including either the weighted FCI or the full list of comorbid conditions predicted more variance than did models containing the additively scored FCI.

Our results are not directly comparable to those of Groll et. al because their sample though largely consisting of patients with spinal impairments and osteoporosis was not categorized as having either lumbar, cervical or other body part impairments. Furthermore, their adjusted models predicting physical function may have contained a different set of case-mix variables (the list of case-mix variables used in their models was not reported). Nevertheless, our findings are largely consistent in that use of a weighted index added little to the explanatory power of their adjusted models.

Contribution of FCI to models containing Intake FS

Consistent with the findings from the crude models, results of our comprehensive risk-adjustment models also showed that that the full list of comorbid conditions and the sub-group specific weighting added greater explanatory power as compared to the additively scored FCI for models of patients in the hip, elbow and neurological group.

There was only one group (shoulder) for which the additively scored FCI had a marginally better explanatory power than sub-group weighted FCI when added to models already containing intake FS. This finding was confirmed in our sensitivity analyses, which could be due to clinical characteristics of the shoulder group that were unobserved in our data. Taken together these findings suggest that the additively scored FCI is appropriate as a risk adjuster for those with cervical, shoulder, lumbar, wrist/hand, knee and foot/ankle impairments and the subgroup specific weighted FCI is more appropriate for those with hip, elbow and neurological impairments, but the differences were small.

Relationship of FCI and weighted FCI

We examined the relationship between the additively scored FCI and sub-group specific weighted FCI scores by creating scatter plots (Figure 1). The scatter plots showed that the relationship between additively score FCI and sub-group specific weighted FCI is not constant across impairment groups. This is not surprising given that the additively scored FCI gives equal value to all comorbidities, whereas the sub-group specific weighted FCI gives higher weighting to specific comorbidities that were found to be most predictive of FS. Further research is needed to understand the clinical relationship of specific comorbidities to FS for each impairment group.

Discussion of the Literature

Our models showed that discharge FS was largely predicted by patient characteristics included in our model, and addition of any of our measures of functional comorbidities added relatively little to models containing other case-mix variables including intake FS for most impairment groups. Overall, these results were not unexpected given that Hart & Connolly, in a similar type of analysis of the FOTO database, found that the additively scored FCI added only 0.018 to the model R^2 value for patients with all types of impairments. However, our findings indicated that use of the weighted FCI or list of condition variables was particularly useful in increasing explanatory value for models predicting discharge FS for the elbow and neurological groups.

Limitations

Our study has several limitations. We corrected for observed differences among those who were and were not censored using inverse probability weighting. Inverse probability weighting involves giving different weights to patients depending on their likelihood of being included in the sample of complete data, which is analogous to using survey weights, where subjects more likely be selected into the study are given less weight in the analysis. Thus, we expect that the impact of missing data was well controlled and any potential bias introduced was negligible. However, it is always possible that there were important unobserved factors that differed across censoring status that we could not adjust for in our inverse probability weighting.

Another limitation to our study is that FS data were collected using body-part specific CAT administrations developed by FOTO. Groll used the PF-10, a different but related measure of functional status in development of the FCI. Each CAT estimates a measure of FS, and each of the CAT administrations has been well validated. Our data did not contain the PF-10 variables or scores, thus we could not replicate our models using the PF-10 as the outcome measure. We are uncertain as to whether or not our results would have been different if using the PF-10. In addition, completion of the FCI and answering the FS items in the CAT depend on patient’s health literacy as well as recall concerning identification of comorbid conditions and the perception of their functional abilities. It is possible that these factors affect patients’ responses to these questions, but we do not know how recall bias may have impacted our results.

The data analyzed came from CAT administrations targeting the ‘primary’ reason for rehabilitation. Approximately 5% of patients in the FOTO database represent people being treated for impairments from two body parts. Because measures of FS from two body parts for which the patient is receiving rehabilitation cannot be considered independent, we only analyzed data from the primary CAT, i.e., the patient and staff member considered the ‘primary’ body part most important for their current rehabilitation episode. We are unaware of the implications of this primary body part emphasis on FS outcomes or relation between FCI and FS.

Conclusion

In conclusion, we found that functional comorbidities contribute to multivariate risk-adjustment models for functional status for patients undergoing outpatient rehabilitation. Our findings suggest that the format for adding functional comorbidities should vary by the impairment group. Use of additive scoring of the FCI as proposed by Groll [22] is most appropriate for patients with shoulder, cervical and lumbar impairments. The use of population-specific weights adds an extra analytical burden and has little utility for these types of patients. However, alternative scoring of the FCI is preferable for patients with knee, foot/ankle, wrist/hand, hip, elbow or neurological impairments. We recommend using either a sub-group specific weighted index similar to the one we developed in this study or, if sample size allows, use of separate comorbid condition variables to maximize explanatory contribution.

What is new?

- Use of the FCI as traditionally scored adds relatively little to predictive ability of already comprehensive risk-adjustment models.
- Use of a weighted index or list of condition variables appears to be of greater value, particularly for patients with hip and neurological conditions.
- This study provides data on the utility of including comorbid conditions into risk-adjustment models for 9 groups of patients undergoing outpatient rehabilitation.
- Many believe that it is time to develop alternative payment systems that reward effectiveness and efficiency of services. However, before we take this step, we need to develop accurate risk-adjustment methods that help produce valid patient comparisons.

Acknowledgments

Funding for this research was provided by the National Institute of Child Health and Human Development (NICHD) R03 HD053780-01A1

References

1. Maxwell, S.; Bsseggio, C.; Storeygard, M. Part B Therapy Services under Medicare, 1998–2000. Health Care Financing Administration; 2001.
2. Dynacorp. Outpatient Therapy Utilization. 2002 [cited 2005 August 9]. Available from: http://www.cms.hhs.gov/providers/therapy/dynacorprpt.asp
3. Chan L. The state-of-the-science: challenges in designing postacute care payment policy. Arch Phys Med Rehabil 2007;88(11):1522–1525. [PubMed: 17964899]
4. Hart, D.; Connolly, J. Pay-for-Performance for Physical Therapy and Occupational Therapy: Medicare Part B Services in Grant #18-P-93066/9-01. Health & Human Services/Centers for Medicare & Medicaid Services; 2006.
5. Porter, ME.; Teisberg, EO. Redefining health care : creating value-based competition on results. Vol. xvii. Boston, Mass: Harvard Business School Press; 2006. p. 506
6. Haley SM, et al. Activity outcome measurement for postacute care. Med Care 2004;42(1 Suppl):I49–I61. [PubMed: 14707755]
7. Guccione AA, et al. Development and testing of a self-report instrument to measure actions: outpatient physical therapy improvement in movement assessment log (OPTIMAL). Phys Ther 2005;85(6):515–530. [PubMed: 15921473]
8. Resnik L, Feng Z, Hart DL. State regulation and the delivery of physical therapy services. Health Serv Res 2006;41(4 Pt 1):1296–1316. [PubMed: 16899008]
9. Hart DL, et al. Computerized adaptive test for patients with foot or ankle impairments produced valid and responsive measures of function. Qual Life Res 2008;17(8):1081–1091. [PubMed: 18709546]
10. Hart DL, et al. A computerized adaptive test for patients with hip impairments produced valid and responsive measures of function. Arch Phys Med Rehabil 2008;89(11):2129–2139. [PubMed: 18996242]
11. Hart DL, et al. Computerized adaptive test for patients with knee impairments produced valid and responsive measures of function. J Clin Epidemiol 2008;61(11):1113–1124. [PubMed: 18619788]
12. Iezzoni, L. Dimensions of Risk, in Risk adjustment for measuring healthcare outcomes. Iezzoni, L., editor. Chicago: Health Administration Press; 1997. p. 43-168.
13. Hofer TP, et al. The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease. Jama 1999;281(22):2098–2105. [PubMed: 10367820]
14. Di Fabio RP, Boissonnault W. Physical therapy and health-related outcomes for patients with common orthopaedic diagnoses. J Orthop Sports Phys Ther 1998;27(3):219–230. [PubMed: 9513868]

15. Jette DU, Jette AM. Physical therapy and health outcomes in patients with knee impairments. Phys Ther 1996;76(11):1178–1187. [PubMed: 8911431]
16. Jette DU, Jette AM. Physical therapy and health outcomes in patients with spinal impairments. Phys Ther 1996;76(9):930–941. discussion 942-5. [PubMed: 8790272]
17. Iezzoni, L. Risk Adjustment for Measuring Health Care Outcomes. Ann Arbor, MI: Health Administration Press; 1994.
18. Greenfield S, et al. Principles and practice of case mix adjustment: applications to end-stage renal disease. Am J Kidney Dis 1994;24(2):298–307. [PubMed: 8048437]
19. Greenfield S, et al. Profiling care provided by different groups of physicians: effects of patient case-mix (bias) and physician-level clustering on quality assessment results. Ann Intern Med 2002;136(2):111–121. [PubMed: 11790062]
20. Charlson ME, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373–383. [PubMed: 3558716]
21. Kaplan MH, Feinstein AR. The importance of classifying initial co-morbidity in evaluatin the outcome of diabetes mellitus. J Chronic Dis 1974;27(7–8):387–404. [PubMed: 4436428]
22. Groll DL, et al. The development of a comorbidity index with physical function as the outcome. J Clin Epidemiol 2005;58(6):595–602. [PubMed: 15878473]
23. Hopman WM, et al. Canadian normative data for the SF-36 health survey. Canadian Multicentre Osteoporosis Study Research Group. Cmaj 2000;163(3):265–271. [PubMed: 10951722]
24. Fanuele JC, et al. The impact of spinal problems on the health status of patients: have we underestimated the effect? Spine 2000;25(12):1509–1514. [PubMed: 10851099]
25. Hart DL, et al. Computerized adaptive test for patients with lumbar spine impairments produced valid and sensitive measures of function. Spine. 2009 In Press.
26. Resnik L, Hart DL. Using clinical outcomes to identify expert physical therapists. Phys Ther 2003;83(11):990–1002. [PubMed: 14577826]
27. Resnik L, et al. Predictors of physical therapy clinic performance in the treatment of patients with low back pain syndromes. Phys Ther 2008;88(9):989–1004. discussion 1004-11. [PubMed: 18689610]
28. Resnik L, et al. Benchmarking physical therapy clinic performance: statistical methods to enhance internal validity when using observational data. Phys Ther 2008;88(9):1078–1087. [PubMed: 18689608]
29. Swinkels IC, et al. Comparing patient characteristics and treatment processes in patients receiving physical therapy in the United States, Israel and the Netherlands: cross sectional analyses of data from three clinical databases. BMC Health Serv Res 2008;8:163. [PubMed: 18667062]
30. Swinkels IC, et al. Clinical databases in physical therapy. Physiother Theory Pract 2007;23(3): 153–167. [PubMed: 17558879]
31. Zimmerman DR. Improving nursing home quality of care through outcomes data: the MDS quality indicators. Int J Geriatr Psychiatry 2003;18(3):250–257. [PubMed: 12642895]
32. Resnik L, Dobrykowski E. Outcomes measurement for patients with low back pain. Orthop Nurs 2005;24(1):14–24. [PubMed: 15722968]
33. Cohen BA, et al. The development of an outcomes management system for acute medical rehabilitation. Am J Med Qual 1997;12(1):28–32. [PubMed: 9116528]
34. Groll DL, et al. Assessment of long-term physical function in acute respiratory distress syndrome (ARDS) patients: comparison of the Charlson Comorbidity Index and the Functional Comorbidity Index. Am J Phys Med Rehabil 2006;85(7):574–581. [PubMed: 16788388]
35. Deyo RA, et al. Outcome measures for low back pain research. A proposal for standardized use. Spine 1998;23(18):2003–2013. [PubMed: 9779535]
36. Hart DL. Assessment of unidimensionality of physical functioning in patients receiving therapy in acute, orthopedic outpatient centers. J Outcome Meas 2000;4(1):413–430. [PubMed: 11272593]
37. Hart DL, Wright BD. Development of an index of physical functional health status in rehabilitation. Arch Phys Med Rehabil 2002;83(5):655–665. [PubMed: 11994805]
38. Linacre JM. Estimating measures with known polytomous item difficulties. Rasch Measurement Transactions 1998;12(2)

39. Hart DL, Mioduski JE, Stratford PW. Simulated computerized adaptive tests for measuring functional status were efficient with good discriminant validity in patients with hip, knee, or foot/ankle impairments. J Clin Epidemiol 2005;58(6):629–638. [PubMed: 15878477]
40. Hart DL, et al. Simulated computerized adaptive test for patients with shoulder impairments was efficient and produced valid measures of function. J Clin Epidemiol 2006;59(3):290–298. [PubMed: 16488360]
41. Hart DL, et al. Simulated computerized adaptive test for patients with lumbar spine impairments was efficient and produced valid measures of function. J Clin Epidemiol 2006;59(9):947–956. [PubMed: 16895818]
42. Ware J Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996;34(3):220–233. [PubMed: 8628042]
43. Ware, JE., Jr, et al. SF-36 Health Survey: Manual and Interpretation Guide. Boston: The Health Institute New England Medical Center; 1993.
44. Andrich D. A rating formulation for ordered response categories. Psychometrika 1978;43(4)
45. Wang YC, et al. Clinical interpretation of computerized adaptive test-generated outcome measures in patients with knee impairments. Arch Phys Med Rehabil 2009;90(8):1340–1348. [PubMed: 19651268]
46. Deutscher D, et al. Implementing an integrated electronic outcomes and electronic health record process to create a foundation for clinical practice improvement. Phys Ther 2008;88(2):270–285. [PubMed: 18042656]
47. Wainer, H., editor. Computerized Adaptive Testing. A Primer. Erlbaum Associates: Lawrence; 2000.
48. Haley SM, et al. Score comparability of short forms and computerized adaptive testing: Simulation study with the activity measure for post-acute care. Arch Phys Med Rehabil 2004;85(4):661–666. [PubMed: 15083444]
49. Hart DL, et al. Computerized adaptive test for patients with knee impairments produced valid and responsive measures of function. J Clin Epidemiol. 2008
50. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology 2000;11(5):550–560. [PubMed: 10955408]
51. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology 2000;11:550–560. [PubMed: 10955408]
52. Lubotsky D, Wittenberg M. Interpretation of regression with multiple proxies. Review of Economics and Statistics 2006;88(3):549–562.
53. StataCorp. Stata Statistical Software: Release 9. StataCorp LP: College Station, TX; 2005.

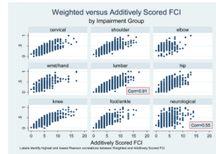
Weighted versus Additively Scored FCI
by Impairment Group
cervical
shoulder
elbow
wrist/hand
lumbar
hip
knee
foot/ankle
neurological
Corr=0.91
Corr=0.55
Additively Scored FCI
Labels identify highest and lowest Pearson correlations between Weighted and Additively Scored FCI

Figure 1.

Table 1

Outcomes and Cohort Characteristics

	Cervical N=5,563	*Shoulder N=8,545*	*Elbow N=1,642*	*Wrist/hand N=2,372*	*Lumbar N=11,109*	*Hip N=3,112*	*Knee N=7,559*	*Foot/ankle N=4,203*	*Neurological N=820*	*Total N=44,925*
	Mean (sd)	**Mean (sd)**	**Mean (sd)**	**Mean (sd)**	**Mean (sd)**	**Mean (sd)**	**Mean (sd)**	**Mean (sd)**	**Mean (sd)**	**Mean (sd)**
Outcomes										
Intake FS	50.24 (12.27)	50.85 (15.57)	46.39 (11.92)	43.65 (14.33)	49.92 (13.65)	49.46 (13.97)	43.15 (14.50)	47.67 (14.05)	46.34 (15.56)	48.23 (14.38)
Discharge FS	64.27 (17.71)	69.58 (15.90)	63.14 (16.89)	61.38 (17.34)	65.52 (16.87)	64.02 (17.81)	63.07 (17.55)	65.48 (17.64)	59.28 (19.57)	65.20 (17.31)
Characteristics										
Age	50.70 (15.00)	53.63 (14.84)	49.10 (13.47)	49.34 (15.84)	51.84 (16.15)	57.38 (16.41)	52.08 (16.18)	49.14 (15.32)	62.14 (16.27)	52.17 (15.80)
No. of Surgeries	0.17 (.55)	0.37 (.64)	0.31 (.68)	0.58 (79)	0.26 (.68)	0.43 (.81)	0.87 (.93)	0.44 (.77)	0.39 (.92)	0.42 (.77)
Weighted FCI[**]	0.23 (.11)	0.22 (.13)	0.33 (.10)	0.23 (.12)	0.15 (.14)	0.22 (.22)	0.22 (.22)	0.16 (.15)	0.34 (.21)	0.21 (.15)
	Median (Q1,Q3)	**Median (Q1,Q3)**	**Median (Q1,Q3)**	**Median (Q1,Q3)**	**Median (Q1,Q3)**	**Median (Q1,Q3)**	**Median (Q1,Q3)**	**Median (Q1,Q3)**	**Median (Q1,Q3)**	**Median (Q1,Q3)**
FCI[*] **(0–18)**	2 (1,3)	2 (1,3)	1 (0,2)	1 (0,3)	2 (1,3)	2 (1,3)	2 (1,3)	1 (0,3)	3 (1,5)	2 (1,3)
Number of Visits	8 (6,12)	10 (6,15)	9 (6,12)	8 (6,12)	8 (6,12)	8 (6,12)	10 (6,15)	9 (6,12)	9 (5,17)	9 (6,13)
	%	%	%	%	%	%	%	%	%	%
Exercise History:										
1–2 x/week	27.5	25.4	28.3	28.6	25.9	22.6	23.7	25.6	16.3	25.4
≥ 3 x /week	37.0	37.5	37.1	33.6	37.5	42.2	39.2	42.9	33.5	38.3
Seldom or never	35.4	37.1	34.7	37.8	36.6	35.2	37.1	31.6	50.1	36.3
Female	66.8	55.8	56.2	62.9	59.1	67.4	60.3	63.3	52.6	60.6
Payer Type:										
Auto Insur.	5.8%	5.3%	4.5%	4.4%	5.0%	5.3%	5.6%	5.1%	6.1%	5.3%
Early Inter.	1.8%	0.3%	0.4%	0.3%	0.5%	0.1%	0.2%	0.3%	0.1%	0.5%
HMO	2.6%	2.6%	2.1%	3.2%	3.6%	2.6%	3.0%	2.5%	6.8%	3.0%
Indemnity	15.6%	19.4%	11.1%	14.8%	20.5%	29.1%	20.5%	14.3%	45.4%	19.5%

	Cervical N=5,563	*Shoulder N=8,545*	*Elbow N=1,642*	*Wrist/hand N=2,372*	*Lumbar N=11,109*	*Hip N=3,112*	*Knee N=7,559*	*Foot/ankle N=4,203*	*Neurological N=820*	*Total N=44,925*
Litigation	1.3%	0.6%	0.6%	1.0%	0.8%	0.5%	0.5%	0.5%	1.6%	0.8%
Medicare A	54.5%	51.8%	51.9%	42.1%	50.3%	50.2%	53.4%	59.1%	36.2%	51.8%
Medicare B	8.2%	13.7%	23.2%	27.9%	12.5%	5.2%	10.6%	11.6%	0.7%	12.3%
Medicaid	10.3%	6.4%	6.2%	6.2%	6.9%	7.0%	6.1%	6.6%	3.0%	6.9%
Onset of condition:										
1–21 days	21.1%	16.5%	19.7%	16.6%	22.7%	15.6%	20.5%	12.8%	22.4%	19.1%
22–90 days	31.1%	32.2%	41.0%	43.3%	26.4%	30.2%	31.2%	40.1%	28.4%	31.9%
91+ days	47.8%	51.3%	39.2%	40.1%	50.9%	54.2%	48.3%	47.1%	49.1%	49.0%

[*] FCI Range 0–18

[**] FCI Range 0–1 Auto Insur= Auto Insurance

FS= Functional Status

Early Inter.= Early Intervention (pediatric)

HMO = Health Maintenance Organization

Medicare A= Hospital Insurance (HI) program, covers inpatient hospital services, skilled nursing facility, home health, and hospice care.

Medicare B= the Supplementary Medical Insurance (SMI) program helps pay for physician, outpatient, home health, and preventive services

Table 2

Prevalence of comorbidities by impairment group

	Cervical N=5563 Percent	Shoulder N=8545 Percent	Elbow N=1642 Percent	Wrist/hand N=2372 Percent	Lumbar N=11,109 Percent	Hip N=3112 Percent	Knee N=7559 Percent	Foot/ankle N=4203 Percent	Neurological N=820 Percent	Total N=44,925 Percent
Arthritis	38	37	24	30	40	48	49	28	40	39
Angina	01	01	01	01	01	01	01	01	04	01
Anxiety	13	09	09	09	11	09	09	09	16	10
Asthma	10	10	09	11	10	10	10	10	09	10
Obesity	29	35	33	36	.37	.33	.46	.41	.33	37
CHF	03	03	02	03	03	04	03	03	08	03
COPD	03	02	01	02	02	03	02	02	04	02
CVA	03	03	02	03	02	04	02	02	30	03
Depression	13	10	09	09	13	11	10	10	16	11
DJD	50	28	21	18	52	30	23	22	26	34
Diabetes	08	12	08	09	10	10	11	10	20	10
GI	16	15	14	12	17	16	14	13	19	15
Hearing loss	07	07	05	05	07	08	06	05	16	07
MI	02	03	02	03	03	04	03	03	08	03
Neurological	02	01	01	01	01	02	01	01	13	02
Osteoporosis	08	08	05	07	09	12	08	07	11	08
PVD	01	01	00	01	01	02	01	01	04	01
Visual problems	14	12	09	08	16	14	10	09	19	12

Table 3

Comparison of R^2 of models predicting discharge functional status (FS) using different methods for scoring the FCI

Group	N	Additive FCI	Weighted FCI	18 separate comorbid conditions
Cervical	5,563	.0416	.0481	.0484
Shoulder	8,545	.0336	.0347	.0347
Elbow	1,642	.0232	.0499	.0499
Wrist/hand	2,372	.0312	.0458	.0458
Lumbar	11,109	.0638	.0720	.0720
Hip	3,112	.0941	.1160	.1160
Knee	7,559	.0682	.0818	.0818
Foot/ankle	4,203	.0583	.0685	.0685
Neurological	820	.0599	.1400	.1400

Table 4

Summary of R^2 of crude and adjusted models predicting discharge functional status (FS) by impairment group

	N	*Intake FS*	*Intake FS + additive FCI*	*Intake FS + weighted FCI*	*Intake FS + 18 separate condition variables*	*Full model* No FCI*	*Full model* + additive FCI*	*Full model* + weighted FCI*	*Full model* + 18 separate condition variables*
Cervical	5,563	.2481	.2572	.2579	.2595	.2866	.2897	.2903	.2912
Shoulder	8,545	.1399	.1558	.1553	.1560	.1738	.1814	.1807	.1818
Elbow	1,642	.1616	.1676	.1821	.1854	.1984	.2013	.2155	.2197
Wrist/hand	2,372	.1555	.1684	.1773	.1791	.1906	.1982	.2051	.2069
Lumbar	11,109	.2321	.2524	.2567	.2581	.3192	.3236	.3249	.3260
Hip	3,112	.2934	.3166	.3257	.3276	.3377	.3466	.3519	.3556
Knee	7,559	.2067	.2317	.2376	.2381	.2535	.2615	.2636	.2647
Foot/ankle	4,203	. .2544	.2739	.2807	.2815	.3088	.3151	.3201	.3223
Neurological	820	.2921	.3047	.3419	.3641	.3521	.3585	.3939	.4033

* Full model controls for age, gender, symptom onset, exercise history, payer type and surgical history.