

Gross DP, Battie MC, Cassidy JD. The Prognostic Value of Functional Capacity Evaluation [FCE] in Patients With Chronic Low Back Pain: Part 1. Timely Return to Work. Spine 2004;29:914-919.

Design: Observational prognostic study

Reviewed, no change to conclusions November 2016

Population/sample size/setting:

- 114 workers (74% male, mean age 41) undergoing FCE at the Alberta Workers' Compensation Board in 1999 and contributing complete data at a one-year follow-up to an exploratory analysis of the predictive value of FCE on return to work (RTW)
- 132 workers (71% male, mean age 40) undergoing FCE in the same setting in 2000 were studied and contributed complete data to a confirmatory analysis of the predictive value of FCE on RTW
- Inclusion criteria were compensated work-related injury with any ICD-9 code indicating low back injury, FCE done at least 6 weeks after the injury, and medical clearance to participate in FCE with no further medical investigation or treatment recommended
- No apparent exclusion criteria; all FCE claimants were eligible for inclusion

Main outcome measures:

- All claimants underwent the Isernhagen Work Systems FCE protocol, with claimant performance on each of 25 FCE tasks given a pass/fail rating for physical job demands
 - o Job demands were based on employer report and direct work site measurement, or on worker self-report when these were not available
 - o In addition, the maximum weight lifted in a floor-to-waist lift task was analyzed in isolation from the other FCE tasks: once as a variable with 10 kg increments, and once with a pass/fail rating
- The original study plan was to classify a worker as safe for RTW only if all 25 FCE tests were passed; however, only 5 workers passed all FCE tests, and the number of successful FCE tasks was substituted as the prognostic factor to be evaluated
- For the 1999 exploratory sample, 68% were receiving TTD benefits at the time of FCE, with a median duration of TTD benefits of 32 days (4 patients received TTD for the entire follow-up year)
- A higher number of failed FCE tasks was related to a longer time receiving TTD benefits; each failed FCE task meant that the claimant was 9% less likely to terminate TTD benefits at any time during the study period
 - o The adjusted FCE model (FCE tasks adjusted for pain index, clinician recommendation following FCE, age, pre-injury salary, and number of health visits preceding FCE) explained 14.8% of the time to suspension of TTD benefits

- The floor-to-waist lift test also was related to the time receiving TTD benefits; each 10 kg increment in weight lifted was related to a 49% greater likelihood that the claimant terminated TTD benefits at any time during the study period
 - o The floor-to-waist lift by itself explained 10.5% of the time to TTD benefits when adjusted for the same variables which were used to adjust the number of FCE tasks
- A confirmatory analysis was carried out in 86 claimants who had FCE in 2000; the predictive properties of the full FCE and the floor-to-waist lift were similar to those measured for the 1999 claimants
- When time to claim closure was analyzed similarly to time on TTD benefits, the predictive performance of FCE and the floor-to-waist lift was similar to the results for TTD benefits

Authors' conclusions:

- It is sometimes recommended that RTW be recommended only when a claimant passes all of the tasks on the FCE
- This recommendation would prevent many claimants from returning to work; only 4% of the claimants achieved the goal of passing all FCE tasks, but nearly all of them closed their claims and terminated TTD benefits during the year following their FCE
- Better performance on FCE (ascertained by fewer failed FCE tasks) did predict faster RTW, but the FCE was a weak predictor of the time required for RTW; the adjusted data for FCE tasks in the 2000 claimant cohort captures only 10.9% of the time to TTD termination
- It is much simpler to measure only the floor-to-waist task of the FCE, and this task alone explained almost as much variation in TTD time as the entire FCE battery of tests
- Most of the variation in time required for RTW is related to factors other than the physical capacity to perform job tasks
- It is reasonable to test only the floor-to-waist lift in low back pain claimants; it is only weakly related to time needed for RTW, but the full FCE battery is also weakly related to time for RTW

Comments:

- The study design captures most of the quality criteria for studies of prognosis: the study population seems to be representative of the population of interest, the measurement of the purported prognostic factor (FCE task performance) is standardized and generally recognized, the outcome is measured within a pertinent time frame which should capture most of the RTW events of interest
- However, there is a very large amount of missing data; only 77 of the 114 claimants in the 1999 cohort and 86 of 132 claimants in the 2000 cohort were analyzed
 - o The authors state that the claimants with missing data were disabled for much longer times than the claimants with complete data (1293 days vs. 553 days), but were similar on other characteristics to those with complete data

- It is reasonable to conjecture that the more chronically disabled claimants had poor RTW compared to the less chronically disabled claimants
- Chronicity was not one of the variables adjusted for in the analysis, but probably should have been included, since it is likely to be associated with poor performance on FCE and with longer time for RTW after the FCE was done
- In Tables 2 and 3, the values for the adjusted PVE (proportion of variation explained) present a problem not discussed by the authors, which may require some modification of their conclusions, but does not totally invalidate them
 - For FCE tests, lifting 10 kg increments, and passing/failing the lifting test, PVE is given for both crude and adjusted analyses
 - In each analysis, the adjusted PVE is less than the crude PVE (e.g., for FCE in the 199 cohort in Table 2, the crude PVE is 24.5%, but the adjusted is only 14.8%)
 - However, the crude and adjusted hazard ratios are nearly identical (0.89 and 0.91 respectively)
 - This suggests that the five variables used to adjust the hazard ratios were not relevant to time for RTW, and added nothing to the analysis
 - Because Cox regression measures the effect of adding new variables in terms of changes in “-2 log likelihood,” it would have been valuable to have had these changes when new variables were added to the crude analysis; from the data in Tables 2 and 3, it appears that these changes were small, and indicated that no new information was being added to the regression analysis
 - Adding irrelevant variables to a multivariable analysis needlessly adds degrees of freedom to the statistical model, and is likely to cause the adjusted PVE to suffer; this is manifested by the fact that adding these variables to the model makes its PVE worse than the crude model, which had only the FCE task failures
 - For the 1999 cohort, therefore, it is more appropriate to say that FCE captures 24.5% of the TTD variation, which is not very good but is better than the 14.8% in the needlessly complex “adjusted” model
 - The PVE for FCE is actually better than for the floor-to-waist lifting task (crude of 24.5% PVE versus 14.1%), but even in this scenario, FCE task performance fails to capture most of the time for RTW in TTD days, and even less of the time for claim closure (only 12.5% of the PVE for the 1999 cohort in Table 3)
 - Therefore, despite what appears to be a flawed regression analysis, the results still support the principle that performance of the physical job demands is only weakly related to RTW and claim closure time

Assessment: Adequate for evidence that (1) FCE task performance is weakly related to time on disability and on time for claim closure, and (2) even claimants who fail on numerous physical performance FCE tasks may be able to return to work

Gross DP, Battie MC. The Prognostic Value of Functional Capacity Evaluation [FCE] in Patients With Chronic Low Back Pain: Part 2. Sustained Recovery. Spine 2004;29:920-924.

Design, population/setting/eligibility criteria as in Part 1 above

Main outcome measures:

- Alberta Workers' Compensation low back claimants who underwent FCE were followed for recurrent back pain events after RTW
- Recurrent event was defined as (1) reopening of the back pain claim, or (2) opening of a new back claim
- Sustained recovery was defined as the avoidance of recurrent events
- Complete data was available for 226 claimants (71% male, mean age 41) from the combined 1999 and 2000 Alberta claimant cohorts
 - o 157 of these were receiving TTD benefits at the time of FCE and had their benefits suspended in the 12 months of observation
 - o 196 claimants had claim closure during the same year; 30 claimants had open claims for the entire year, but were still at risk for recurrent TTD episodes
- Of the 157 claimants whose TTD benefits were terminated, 19 (12%) restarted TTD benefits within the same year
- Of the 196 patients whose claims were closed, 28 (14%) had their claim reopened or filed a new back claim within the first year
- The definition of "recurrent event" was met by 46 of the 226 claimants
- There were 25 tasks in the FCE battery; the median number of failed tasks was 8, and this was used as a cutoff for the first estimation of whether FCE predicted recurrent events
 - o FCE did not predict recurrent events when analyzed in this manner
 - o 16% of those who failed more FCE tasks had recurrent events, but 25% of those with fewer FCE task failures had recurrent events, the opposite of what would be expected if FCE failures predicted recurrent events
- In addition, logistic regression modeling was done with the FCE and recurrent event data
 - o Failing more FCE tasks had an adjusted odds ratio of 0.94 (95% confidence interval, 0.87 to 1.02) for recurrent events
 - o This is a statistically nonsignificant trend towards having 6% fewer recurrent events for every FCE task failed; again, it is inconsistent with the hypothesis that FCE task failures are prognostic for recurrences of back pain after RTW
 - o Adjusting for time to initial TTD suspension and claim closure did not alter the direction or magnitude of the results
- As with Part 1, floor-to-waist testing was examined in a logistic regression model, and was not associated significantly with risk of recurrence

Authors' conclusions:

- Contrary to expectations, better performance on FCE tasks was associated with higher risk of recurrence
- There were some limitations to the study in its reliance on administrative data bases rather than patient self-report, but the validity of FCE as a tool to identify claimants who are safe for RTW is not supported by the data

Comments:

- The analysis of recurrence data by logistic regression is more straightforward than for the time to claim closure by Cox regression, since it does not entail attempts to report proportion of variation explained by the model
- As with the Cox regression in Part 1, the addition of variables (number of previous back claims, pain disability index, gender, and clinician recommendation) did not appreciably influence most estimates of FCE as a prognostic tool
- There were many claimants with missing data, as was the case with Part 1, where case chronicity was associated with having missing data and omission from the analysis; there is no apparent reason to suppose that FCE would perform better with more chronic cases

Assessment: Adequate for evidence that FCE fails to predict which injured workers are likely to have a sustained return to work after resolution of a back claim