

Title: Validation of the Pyruvate Kinase Deficiency Impact Assessment (PKDIA): a patient-reported outcome measure for pyruvate kinase (PK) deficiency

Authors: Shayna Egan, Erin Zagadailov, Adrian Jewett, Junlong Li, Feng Tai, Audra Boscoe, David A. Andrae

Introduction

Pyruvate kinase (PK) deficiency is a rare hereditary disorder characterized by chronic hemolytic anemia. There are no patient-reported outcome (PRO) instruments specific for this condition. The PK Deficiency Impact Assessment (PKDIA), a disease-specific patient-reported outcome (PRO) instrument, was designed as a weekly measure to assess impacts of PK deficiency. The PKDIA was administered as a 12-item assessment in which patients (pts) were asked to rate the frequency of occurrence or difficulty with various experiences or activities using an 11-point numeric rating scale (NRS) ranging from 0-10 or a 5-point verbal descriptor scale (VDS). The instrument underwent psychometric validation using blinded data from the 80 pts participating in the ACTIVATE trial, a randomized, multicenter, double-blind, phase 3 trial of mitapivat, an allosteric activator of PK, in non-regularly transfused adults with PK deficiency (NCT03548220).

Methods

Completion rates and baseline response distributions were characterized using descriptive statistics. Inter-item correlations were estimated and item response theory (IRT) modeling was applied. A scoring system was established with item weighting informed by IRT modeling. The resultant baseline scores were used to assess reliability (internal consistency and test-retest) and validity (convergent and known-groups). Change scores for group comparisons were evaluated via anchor-based methods using the Patient Global Impression of Severity (PGIS) changes from baseline to weeks 12 and 24. The PGIS had a 5-point VDS (Not at all, A little, Moderately, A lot, and Very much).

Results

Baseline PKDIA data were available for 78 pts (97.5% completion rate). Response distributions skewed right with sparse endorsements at the higher levels, especially for NRS-rated items. To facilitate IRT modeling, responses were recoded to a 0-4 scale. Baseline IRT modeling revealed four items were less relevant to the ACTIVATE trial population or did not contribute unique information due to skewness or redundancy. After removing these four items, the final model resulted in excellent fit [Root Mean Squared Error of Approximation (RMSEA) 0.04, 95% CI = (<.01, 0.12)]. The eight retained items assess disease impact on starting things you wanted to get done, daily activities, relationships with friends or family, concentration, physical activity, and need for additional rest or sleep. Sum scores were calculated and converted to T-scores (mean of 50 and standard deviation of 10) employing a scoring conversion (Table). The conversion of sum scores to T-scores is based on the expected a-priori scores (Z) from the IRT model. A higher score represents more severe disease impacts and higher disease burden.

Internal consistency was assessed with McDonald's coefficient, $\omega = 0.90$, indicating a high level of reliability. Test-retest reliability via a two-way mixed effects Intra-Class Correlation Coefficient (ICC) was high (0.87). Convergent validity was established when correlating PKDIA scores with the FACT-An ($|r| = 0.82$), the SF-12 PCS ($|r| = 0.64$), and the SF-12 MCS ($|r| = 0.50$). Known-groups validity was moderate, with mean scores ordering approximately as expected across baseline PGIS ratings and a linear trend ($\eta^2 = 0.51$).

Changes in scores from baseline to weeks 12 and 24 were assessed by stratifying PKDIA scores by changes in PGIS (no change as the reference group). By comparing pts who reported a 2-point worsening in severity

to those who reported no change on the PGIS, estimates of a change in PKDIA scores at weeks 12 and 24 ranged from 2.0 to 13.9 points. Per anchor-based methods, a change of approximately 6 to 10 points could be considered a meaningful change in PKDIA score.

Conclusions

The PKDIA is now the first validated disease-specific PRO to assess disease impacts in pts with PK deficiency. The PKDIA showed high internal consistency, test-retest reliability, and convergent validity. Further research investigating meaningful change scores would be beneficial. Due to the rare nature of PK deficiency, the ACTIVATE study enrolled a small number of pts overall which was a limitation in the anchor-based determination of meaningful change. Thus, the recommended estimates of meaningful change, although interpretable for the current study, are not definitive and could benefit from further investigation.

Table 1. PKDIA scoring table

Sum Score	Z Score	T Score
0	-1.9263	30
1	-1.4949	35
2	-1.2576	37
3	-1.0708	39
4	-0.9096	40
5	-0.7675	42
6	-0.6374	43
7	-0.5181	44
8	-0.4087	45
9	-0.3072	46
10	-0.2111	47
11	-0.1175	48
12	-0.0250	49
13	0.0670	50
14	0.1579	51
15	0.2475	52
16	0.3366	53
17	0.4266	54
18	0.5188	55
19	0.6140	56
20	0.7117	57
21	0.8110	58
22	0.9115	59
23	1.0147	60
24	1.1234	61
25	1.2401	62
26	1.3648	63
27	1.4967	64
28	1.6378	66
29	1.7964	67
30	1.9878	69
31	2.2356	72

Sum Score	Z Score	T Score
32	2.6016	76