Evaluating depression in those who cannot speak for themselves: The utility of the PHQ-9 OV

Rebecca Ferrini, MD, MPH, CMD; Robert Gibson, PhD, JD; Jeff Klein, PhD.

Edgemoor DP SNF, Santee California

INTRODUCTION

The MDS version 3.0 incorporates a more sensitive and specific depression inventory called the PHQ-9 and PHQ-9 Observational Version (OV). Very little has been published about the utility of the PHQ-9 OV. Our facility is a 192-bed not for profit, government-run, fully Medicaid distinct part home with average age 56 yrs.

METHODS

We examined historical data from 10/1/12 to 12/31/12 with PHQ-9 scores, depression diagnoses by a MD/PhD, and individual items for the PHQ-9 OV obtained by Master's level social workers. T tests compared mean PHQ-9 scores of the self and staff report. We compared our score distribution with the MDS 3.0 Evaluation and Validation sample (Saliba et al. 2012). Chi squared statistics compared data.

RESULTS

- Among 191 residents, 61 (32%) required staff assessment (PHQ-9 OV), and 130 (68%) completed the PHQ-9.
- Mean PHQ-9 values for self-assessments were more than four times higher than staff assessments (self 2.7, staff 0.80, T test p<0.001).
- 95% of PHQ-9 OV scores were less than 5 compared with 78% of PHQ-9 scores. Scores above 10 were very rare (7% in PHQ-9 and 2% in PHQ-9 OV).
- When compared with the MDS validation study, we had a higher proportion of low scores and a different distribution (p<0.0001) for both the PHQ-9 and PHQ-9 OV.

ANALYSIS

Clinically significant values were less likely on the PHQ-9 OV compared to the PHQ-9 (p<0.001).
- 20%(12/61) of those who completed PHQ-9 OV had depression diagnosis compared to 35% (41/130) of those completing the PHQ-9 (p=0.087).
- 100% of those diagnosed with depression, screened with the PHQ-9 OV were on antidepressant medications, compared to 80% (37/46) of those who could communicate (p=0.0955)

A "gold standard" for screening or diagnosing depression in those with severe communication difficulties has not been established. The PHQ-9 OV yielded significantly lower mean scores and fewer high scores than the PHQ-9 in both our study and in the validation study. "Irritability" added to the PHQ-9 OV version was the most common symptom noted by staff.

Why might PHQ-9 OV scores be lower than PHQ-9 even though the PHQ-9 OV has an extra item?
- A lower rate of depression or symptoms among those who can’t communicate.
- Staff not using instrument correctly.

ACKNOWLEDGEMENTS

Acknowledgement to Nael Aniloua for her contribution in designing the poster.

ANALYSIS (Continued)

Instrument is less robust when removing subjective aspects of each question focusing only on observational aspects.
- Diagnostic criteria for those without ability to communicate may differ from DSM standard.
- Observers accustomed to seeing long-stay debilitated individuals, don’t identify chronically observed characteristics as notable.
- Staff knowing this is a depression inventory may choose responses to support their beliefs.
- Some PHQ-9 items when used on the OV may be confusing or irrelevant. For instance, how can appetite be assessed in a resident who is on tube feeds?
- Instrument not relevant to vegetative, but perhaps relevant to nearly vegetative individuals either?
- Our sample found two items never selected in our sample of 61 residents (items about feeling bad about self or letting others down or thoughts of self harm).

If the PHQ-9 OV is not capturing depression, how do we screen for depression in those who cannot speak?

Identification of behavioral signs used by clinicians to diagnose and monitor treatments for depression. In non-verbal residents, we use tearfulness and crying (more than half), poor sleep, delusional statements, irritability and aggression.
- Identification of which items in the PHQ-9 OV are the most often associated with depression and response to antidepressants.

REFERENCES

Kroenke, K, Spitzer, RS, Williams, JBW. The PHQ-9
Sellers, D. Afilalo S, O’Cormac Mat, Testing the PHQ-9 Interview and Observational Version (PHQ-9-OV) for MDS 3.0.

CONCLUSIONS

We found lower scores on the PHQ-9 OV compared to the PHQ-9, with a paucity of scores over 10 which are the most likely to be clinically relevant. Certain elements of the PHQ-9 OV may be problematic (appetite, internal feelings, concentration, mobility or motivation). The PHQ-9 OV seems to lack "face validity" with inclusion of items requiring expression of internal emotional states from persons who by definition are unable to do so.

Further research might investigate the correlation between the PHQ-9 and PHQ-9 OV in individuals with very minimal communication abilities and severe disabilities, whether a different cut off should be selected, how to deal with somatic confounders, and whether all the PHQ-9 OV items are necessary or helpful. We propose an instrument to screen for depression in those who cannot speak that includes items such as resistive to care, crying, irritability, poor sleep, aggression and “response to antidepressant therapy.”