

Regional and Temporal Differences in Use of Doctorate Level MADRS Raters in Bipolar Depression Trials

David Daniel, MD¹, John Bartko, PhD², Norman Sartorius, MD, PhD, FRCPsych³, Eduard Vieta, MD⁴, Adam Butler⁵

¹United BioSource Corporation, McLean, VA, ²Independent Consultant, Newville, PA, ³Geneva University, Geneva, Switzerland, ⁴University of Barcelona, Barcelona, Spain, ⁵United BioSource Corporation, Wayne, PA

INTRODUCTION

In multi-country clinical trials, agreement among raters may be complicated by variations in rating practices. The Montgomery-Asberg Depression Rating Scale (MADRS) is a widely recognized and utilized measure in bipolar depression clinical trials. There is no universal agreement regarding the level of professional credentials required to rate the MADRS. As such, we hypothesized that there would be significant variation by region and over time in the use of doctorate vs. non-doctorate raters for the MADRS.

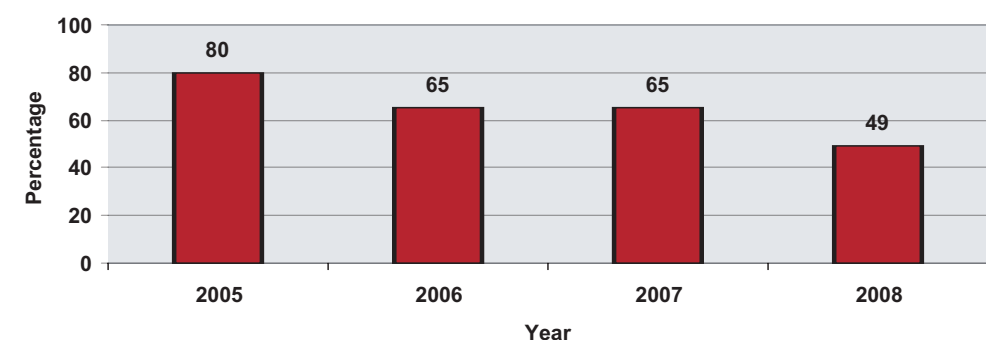
METHODS

This analysis was designed to compare the use of doctorate vs. non-doctorate raters to rate the MADRS across nine industry sponsored multi-center bipolar depression clinical trials over several years (2005, 2006, 2007 and 2008) and among regions [North America (NA) (n=379), South and Central America (SCA) (n=89), Western Europe (WE) (n=65), Eastern Europe (EE) (n=141) and Asia (n=73). The proportion of doctorate level raters was compared by region and by year of certification to rate using the chi-square statistic.

RESULTS

SUMMARY OF RESULTS I: For the five regions combined the differences among the years 2005, 2006, 2007 and 2008 in the use of doctorate vs. non-doctorate MADRS raters was statistically significant ($\chi^2 = 161$, $p < 0.000001$; $df = 4$).

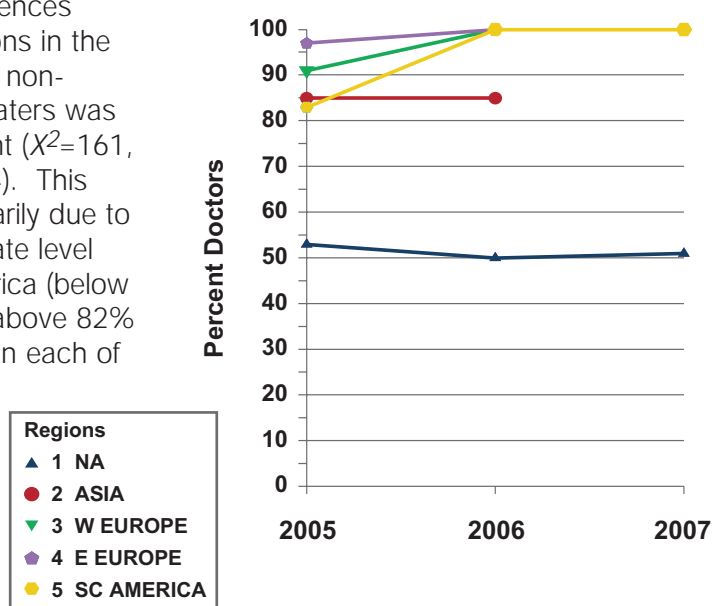
GRAPH I: PERCENTAGE OF ALL RATERS HOLDING A DOCTORATE DEGREE (N = 747)



RESULTS (cont.)

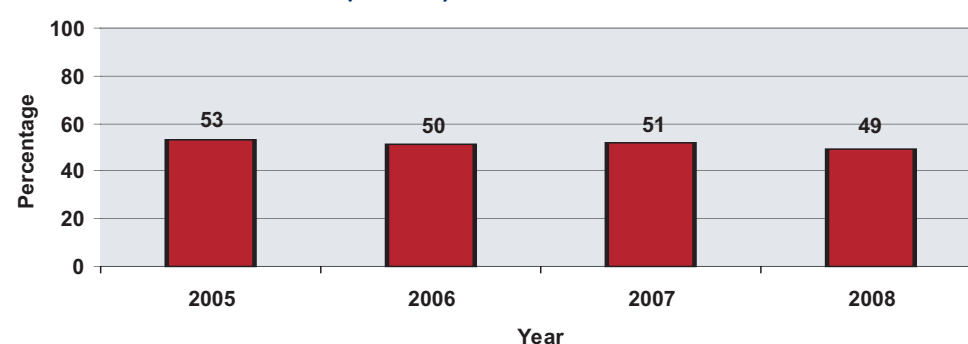
SUMMARY OF RESULTS II: Over the past four years combined, the differences among the five regions in the use of doctorate vs. non-doctorate MADRS raters was statistically significant ($\chi^2=161$, $p < 0.000001$; $df = 4$). This difference was primarily due to lower use of doctorate level raters in North America (below 54%) compared to above 82% for all other regions in each of the years assessed.

GRAPH II: PERCENT DOCTORATE RATERS BY REGION AND YEAR

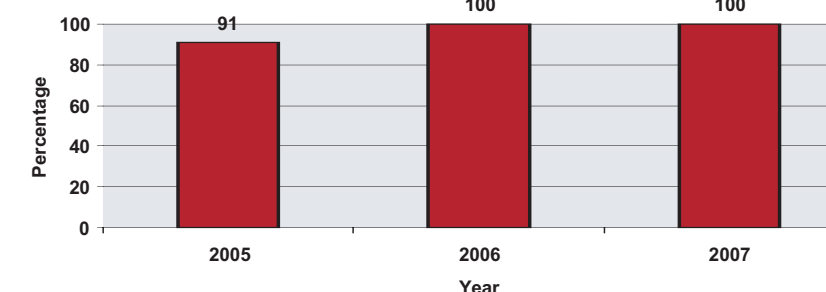


SUMMARY OF RESULTS III: As shown in Graphs III-VII the percentage of raters holding doctorates in North America ($\chi^2 = 0.42$, $p < 0.94$; $df = 3$), Western Europe ($\chi^2 = 2.2$, $p < 0.34$; $df = 2$), Eastern Europe ($\chi^2 = 0.84$, $p < 0.66$; $df = 2$) and Asia ($\chi^2 = 0.001$, $p < 0.97$; $df = 1$) was not statistically significantly changed over time. As shown in each graph below data were not available for all regions for each year. In Central and South America there was a modest increase in the percentage of raters holding doctorates ($\chi^2 = 7.12$, $p < 0.03$; $df = 2$) from 2005-2007.

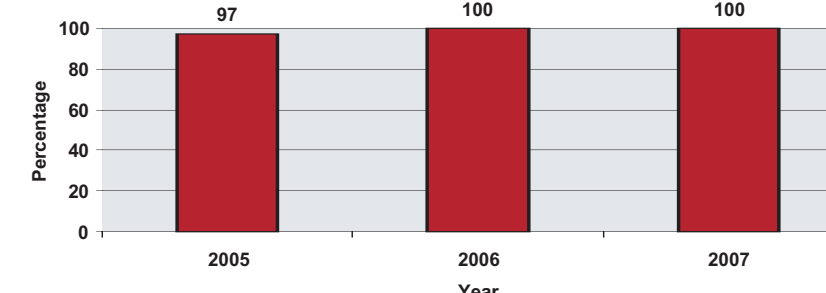
GRAPH III: PERCENTAGE OF NORTH AMERICAN RATERS HOLDING DOCTORATES BY YEAR (n = 379)



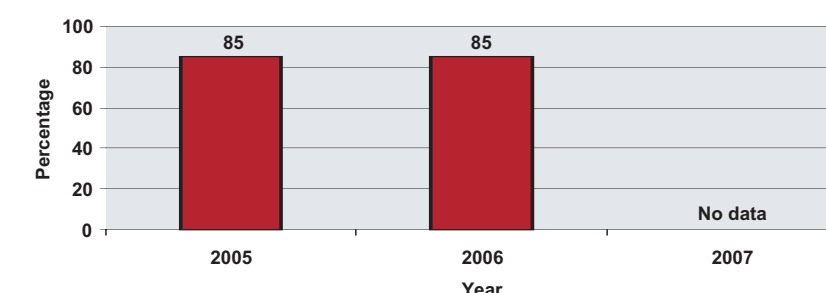
GRAPH IV: PERCENTAGE OF WESTERN EUROPEAN RATERS HOLDING DOCTORATES BY YEAR (n = 65)



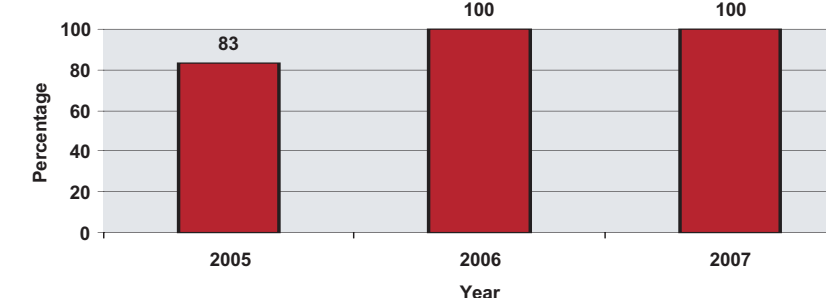
GRAPH V: PERCENTAGE OF EASTERN EUROPEAN RATERS HOLDING DOCTORATES BY YEAR (n = 141)



GRAPH VI: PERCENTAGE OF ASIAN RATERS HOLDING DOCTORATES BY YEAR (n = 73)



GRAPH VII: PERCENTAGE OF CENTRAL AND SOUTH AMERICAN RATERS HOLDING DOCTORATES BY YEAR (n = 89)



CONCLUSION

The results are consistent with significant regional variation in the educational level of MADRS raters. North America utilized a significantly lower percentage of doctorate level raters than the rest of the world. The proportion of raters holding doctorates in North America over time (2005-2008) was not significantly statistically changed and never exceeded 54%. Outside of North America the proportion of raters holding doctorates exceeded 82% in all years evaluated and was stable over time except in Central and South America where there was a modest increase.

Although the ratio of doctorate to non-doctorate MADRS raters appears to have declined over the last four years in the combined sample of all five regions, interpretation must be caveated by the disproportionate number of raters from North America in the sample in 2008. In A longer observation period and more balanced sample geographically might reveal clearer temporal trends. In addition, all raters were trained by the same company which may be another source of sample bias.

Recent publications have focused on diminution of placebo drug separation in a variety of therapeutic areas including depression (Yang, Cusin and Fava, 2005). One recent publication noted an example of higher placebo response and lower drug response in the United States compared to other regions (Kemp, Schooler and Kalali, 2008). The relationship, if any between regional differences in data quality and educational levels of raters is largely unexplored. Future presentations will explore the relationship among educational level of raters, measures of performance in training and results of data surveillance.

REFERENCES

Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979; 134:387-389

Yang H, Cusin C, Fava M: Is There a Placebo Problem in Clinical Trials? *Current Topics in Medicinal Chemistry* 2005; 1077-1086.

Kemp AS, Schooler NR, Kalali AH et al: What is Causing the Reduced Drug-Placebo Difference in Recent Schizophrenia Clinical Trials and What Can be Done About It? *Schizophrenia Bulletin* August 22, 2008 (Epub Ahead of Print)

ACKNOWLEDGEMENT

The authors gratefully acknowledge Gwyneth Moya, MPH for her expert assistance in the preparation and editing of the presentation.

This research was funded by United BioSource Corporation.

Presented at The International Society for CNS Clinical Trials and Methodology (ISCTM) 5th Annual Scientific Meeting 2009