



**Supplementary Figure S7: Illustration of the effect of dataset composition on the prognostic value of biochemical/IHC based PgR status**

Kaplan-Meier analysis was performed in two arbitrarily defined dataset pools: The 882 ER positive patients from the seven datasets in which a prognostic value of OPG was detected (see Supplementary Table S3) were combined as dataset-pool #1. All 1059 ER positive patients from the 11 datasets in which no prognostic value of OPG has been detected were combined as dataset-pool #2.

For 335 of the 882 patients from dataset-pool #1 and 606 of the 1059 patients from dataset-pool #2 information on PgR status from biochemical/IHC assay was available. The Kaplan-Meier graphs display the differences of event free survival according to PgR status as determined by biochemical/IHC assay separately for dataset-pool #1 in (A) and dataset-pool #2 in (B) for exploratory purpose.