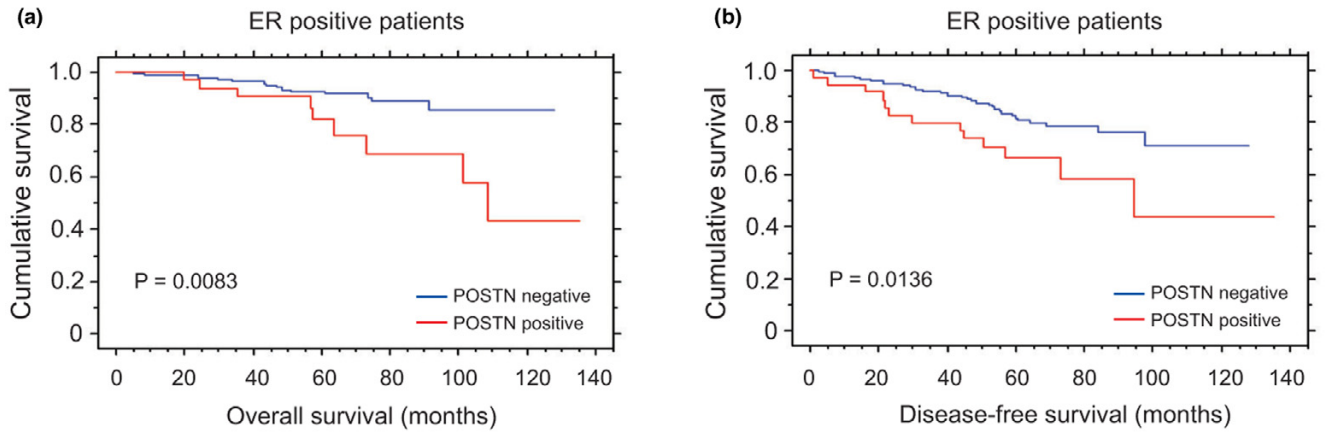
Figure 6



Cumulative Kaplan-Meier curves for epithelial expression of periostin (POSTN). A cohort of poor-prognosis estrogen receptor (ER)-positive tumours was analysed showing: **(a)** a significantly shorter overall survival ($P = 0.0083$); **(b)** a shorter disease free survival ($P = 0.0136$).

malignant samples by normal or reactive stromal cells, as well as the induction of inflammatory genes due to *in vitro* manipulation, could result in false positives. However, verification of the probable epithelial origin of differentially expressed genes can be obtained by comparing expression data from breast epithelial cell lines [22], breast tumour cell lines or, as in the present study, by immunohistochemistry, all of which show that, for example, IL8, is a *bona fide* epithelial tumour-associated product [43,44]. One of the features of normal luminal epithelial cultures is the loss of estrogen receptor expression [45]. The microarray gene expression profiling currently used to classify breast cancers supports the paradigm that ER status is the most important phenotype in breast cancer and has led to the classification of breast cancers into luminal A (ER-positive good prognosis) and luminal B (ER-positive poor prognosis), and ER-negative myoepithelial/basal and HER2 subtypes, each with distinct differences in prognosis and response to therapy [4,5,46]. Genes identified in this study representing the normal luminal epithelial phenotype are distinct from the subset of genes that are associated with ER expression and are used to classify 'luminal' breast tumours. Thus, we are able to define the luminal phenotype independently of ER status. In contrast, our myoepithelial signature contains several members of the previously reported gene clusters identifying basal-like breast cancers. Some of these have been previously identified as myoepithelial genes in the normal breast epithelium, for example, *TIMP3*, *SPARC*, *JAG1*, *PRSS11* and *CAV-1* [11], and some of them, such as *S100A7*, *SPARC* and *CNN1*, have previously been shown individually to be correlated to poor outcome [5,11,47]. Since our cell type specific gene signatures were derived from phenotypically well characterised cell types compared to empirical stratification based on expression data, we were also able to identify a range of myoepithelial type genes in ER-positive tumours as well as those in basal-like breast cancers. Thus, although the majority of the primary breast tumours within our malignant pool were ER-positive 'luminal' tumours, a significant number of up-regulated gene sets also showed myoepithelial expression. The observation of myoepithelial genes such as *SFRP2*, *DCN*, *POSTN*, *LUM*, *COL1A2* and *COL11A1*, which showed higher expression in ER-positive compared to ER-negative breast tumours in two other breast cancer tumour profiling studies [48,49], proved the value of such an approach and demonstrated the heterogeneity of breast tumours with respect to the levels of luminal epithelial and myoepithelial gene expression. The potential clinical significance of the expression of myoepithelial/basal genes in ER-positive tumours has been highlighted by recent data showing that the promoter DNA methylation of the classic myoepithelial marker *S100A2* is correlated with a poor prognosis in ER-positive tumours [50]. In contrast, increased levels of expression of phosphoserine aminotransferase (encoded by *PSAT1*), which was another gene also identified in our myoepithelial transcriptome, was the strongest predictive marker for a poor response to tamoxifen therapy in ER-positive tumours [50]. Our observation that the malignant epithelial expression of POSTN, also a myoepithelial/basal gene, is associated with poorer survival ($P = 0.0083$) in ER-positive tumours demonstrates that the normal epithelial annotation of tumour transcripts can identify many other types of myoepithelial/basal genes, including those associated with a poor outcome.

An important question is whether the expression of myoepithelial/basal genes in breast cancers are responsible for the prognosis and poor response to therapy or are merely surrogate markers thereof. There are several lines of evidence to suggest that POSTN may play a role in the biology of breast cancer [51,52]. POSTN is a ligand of $\alpha_v\beta_3$ integrins and promotes adhesion and migration of epithelial cells [51]. Clinical studies of periostin expression in human cancers have demonstrated