# S2 Text. Localizations.

To further stratify patients by the degree of localization of their active joints, we conducted a bootstrap analysis. In this analysis, we calculated a threshold $x\_{localized}$ based on the proportions of active joints $p$ that are also key joints in high-level factors underlying patient group assignments. For example, if $x\_{localized}=1$, a patient with only knee arthritis would have 100% of their active joints appearing in **<F knees>** and therefore would be *localized* with respect to **<F>** (see Results). We considered patients with $p\geq x\_{localized}$ as having *localized* involvement. The number of such patients is $n\_{x\_{localized}}$. We chose the highest threshold $x\_{localized}\in (0, 0.1, …, 1)$ that satisfied

$$μ\left(n\_{x\_{localized}}\right)<μ\left(n\_{x\_{localized}=1}\right)-SEM(n\_{x\_{localized}=1})$$

where $μ(…)$ is the bootstrap mean and $SEM\left(…\right)$ is the standard error of the mean.

We defined an additional category, *partially localized* involvement, by determining which threshold $x\_{partial}$ had the lowest slope $m$ as calculated by $m=\frac{p\_{x+0.1}-p\_{x-0.1}}{0.2}$. We classified patients with $p<x\_{partial}$ as having *extended* involvement, involvement of joints beyond those defining their underlying high-level factors.