## SUPPLEMENTARY MATERIAL

Studies Identified Through the Systematic Review

Studies identified in the systematic review, that passed Grades of Recommendation, Assessment, Development and Evaluation (GRADE) assessment, or were primary regulatory trials, that had been included in previous technology appraisals of other treatments, were included in the meta-analysis. The systematic review identified 472 studies which reported efficacy data.

Fig. S1 PRISMA chart for the systematic review. *PRISMA* Preferred Reporting Items for Systematic Review and Meta-Analyses.
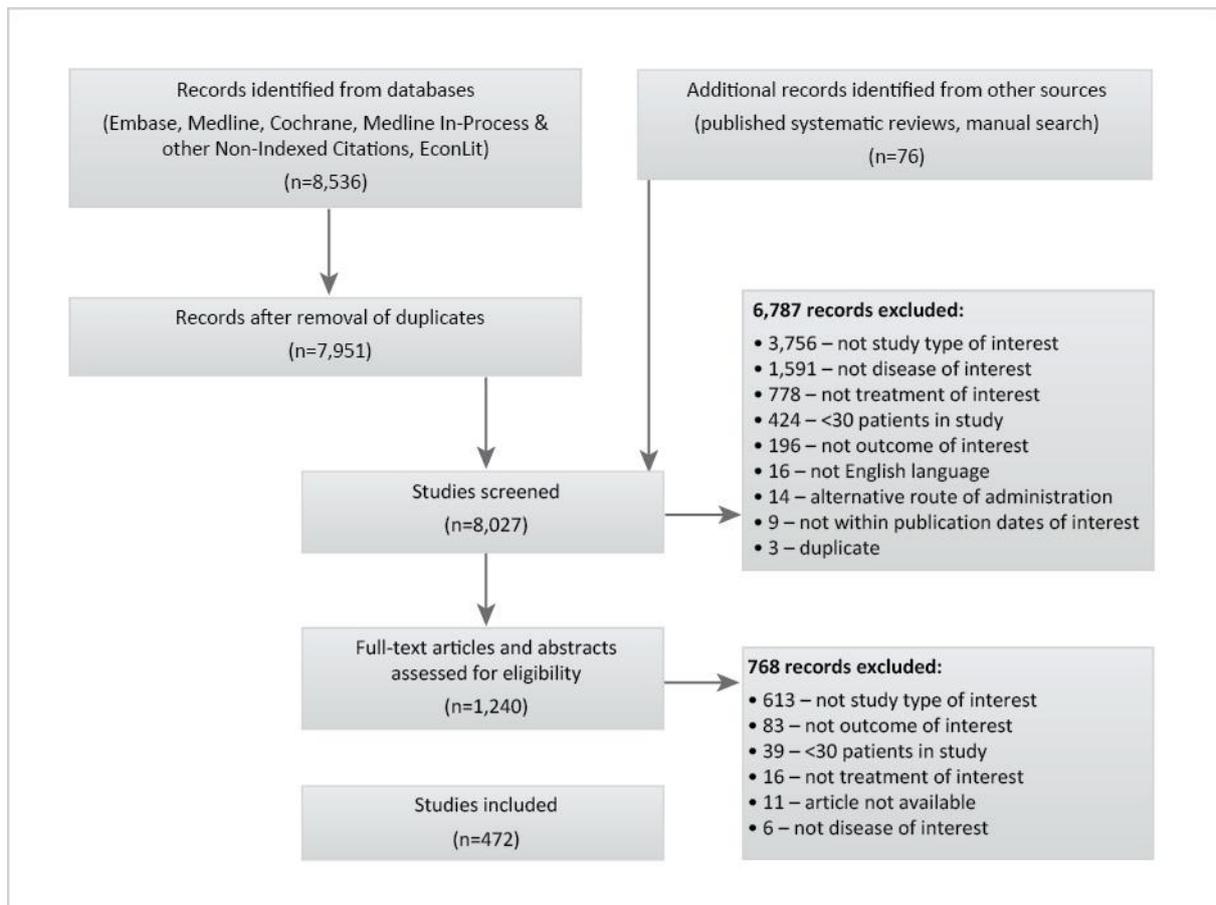
From the 472 efficacy papers that were included in the systematic review, 368 were excluded before GRADE assessment because they reported on trials of agents that were not comparators of interest. In the case of combination arms, both treatments had to be a comparator of interest for the study to qualify for GRADE assessment.

In total, 104 studies were assessed for bias and study quality using the GRADE assessment. The large majority of studies ($n = 96$; 92%) were assessed as very low quality or low quality, either because they were not randomized ($n = 65$) or because criterion 1 requiring adequate sequence generation was not satisfied ($n = 31$). A further two studies were assessed as low quality because, although criterion 1 requiring adequate sequence generation was met, one or more of criteria 2–7 were not.

Six studies passed GRADE assessment: two were high-quality studies in which all criteria were satisfied and five were assessed as moderate quality. Studies rated as moderate quality had to satisfy criteria 1–3 requiring studies to have adequate sequence generation, allocation concealment, and blinding. Open-label studies did not pass criteria 2 or 3 and therefore could not be rated higher than low quality in the GRADE assessment, excluding them from the meta-analysis.

Fig. S2 Study selection using GRADE criteria. *GRADE* Grades of Recommendation, Assessment, Development and Evaluation.