**Glossary**

**Comparative effectiveness-** The goal of comparative effectiveness research is to provide information to help make informed decisions about which treatment option is most appropriate in a given situation. Randomised controlled trials (RCTs) and real-world data can both be used to address questions around comparative effectiveness, however there is a greater potential for bias and confounding with RWD (see *propensity score and* *unmeasured confounding*)

**ECOG**- Eastern Cooperative Oncology Group (ECOG) performance status is a measure of a cancer patient's physical functioning and ability to carry out daily activities. A higher score on the scale indicates a greater level of physical impairment or disability.

**External/Synthetic Control arm-** Using existing patient level data either from a historical trial or real-world data, a control group is created that is designed to be similar to a group of patients in a single-arm trial. This control arm can then be compared to the single-arm trial to provide *comparative effectiveness* results for an outcome.

**Missing data-** Real-world data are often missing for intentional (e.g., a patient does not need certain laboratory tests) or unintentional (e.g., lack of routine follow-up) reasons, and this missingness can result in bias. Within real-world studies, various imputation models have been successfully applied and these include mean substitution, as well as advanced statistical methods, such as multiple imputation.

**Overall survival-** measure of the length of time that a person with cancer lives after a certain time point (e.g., diagnosis or treatment initiation)

**Progression-free survival-** measure of the length of time that a person with cancer lives without their disease worsening after a certain time point (eg diagnosis or treatment initiation)

**Propensity score-** Treatment groups in the real-world can differ in characteristics that are associated with the outcome of interest. These differences can occur because a characteristic is part of the treatment decision (and not a random choice like an RCT), an issue termed confounding by indication. Measured confounders can be addressed (“statistically adjusted” for) through multivariate and propensity score analysis. Researchers have traditionally relied upon multivariable regression models to adjust for differences in patient characteristics, however, they are now turning to propensity score and inverse probability of treatment weighting (IPTW) because of the stronger theoretical and statistical basis for these methods. Propensity scores are estimated using regression-based methods; the score reflects the probability of treatment assignment. Weights are assigned to patients based on the inverse of their propensity score. IPTW results in a pseudo-population where weights determine the extent to which each patient contributes. This pseudo-population can then be compared, for example, comparing the risk of the study main outcome with hazard ratios from Cox proportional hazards regression models.

**Quantitative bias analysis-** methods that have been developed to assess the impact of confounding and other sources of systematic error on study results.

**Real-world data**- data collected from patients outside from a controlled (trial) setting. Can include electronic medical records or insurance claims databases.

**Time to treatment discontinuation-** measure of the length of time a patient is treated with a given treatment, regardless of the reason for discontinuation (e.g., lack of treatment response or treatment induced adverse reaction). Skipping treatment cycles is not considered treatment discontinuation if the treatment is resumed for subsequent cycles.

**Unmeasured confounding-** the potential influence of factors that are not accounted for in a statistical analysis (see *propensity score*), but which may be related to both the exposure being studied and the outcome. This issue is not so relevant for powered RCTs where randomisation takes care of both measured and unmeasured confounders, but it is an issue that can potentially afflict all RWD studies.