Third party subjective assessment of aesthetic outcome following breast conserving therapy (BCT) is often meant to be objective. We analysed intra- and inter-rater agreement of subjective third party assessment and agreement with a semi-automated objective software evaluation tool (BCCT.core).

We presented standardized photographs of 50 patients, taken shortly and one year after surgery to a panel of five breast surgeons, six breast nurses, seven members of a breast cancer support group, five medical and seven non-medical students. In two turns they rated aesthetic outcome on a four point scale. Moreover the same photographs were evaluated by the BCCT.core software, a semi-automated software to measure aesthetic outcome objectively.

Agreement within the four users of the semi-automated BCCT.core software was very good (multiple kappa = 0.8). Intra-rater agreement in the panel members was moderate to substantial (k = 0.4-0.5; wk = 0.6-0.7; according to different subgroups and times of assessment). In contrast inter-rater agreement was only slight to fair (multiple kappa = 0.1-0.3). Agreement between the panel participants and the software was fair (weighted kappa = 0.24 to 0.45). There were neither covariates explaining better or worse agreement between panel rating and software nor relevant differences between the two times of assessment.

Subjective third party assessment only fairly agree with objective BCCT.core evaluation just as third party participants do not agree well among each other.