To increase understanding of potential variability in PFS measurement, relevant past studies of the concordance between INV- and ICR-assessed PFS across clinical studies concluded that, while there is variability between individual INV and ICR results, overall there is strong concordance and no evidence of systematic bias.1–4 Moreover, more recent individual clinical studies have reported large differences between INV- and ICR-assessed PFS; for example, the SOLO-2 study, which compared olaparib versus placebo as maintenance treatment for platinum-sensitive ovarian cancer, reported an INV-assessed median PFS of 25.9 months and an ICR-assessed median PFS of 17.0 months.5,6 As such, assessing PFS should be an integral part of potential sources of variability in PFS measurement.

### Methods

We searched PubMed using the same criteria; full articles were reviewed and studies of solid tumors, hematologic malignancies, and other cancers were included.8 For non-comparative trials, we calculated the absolute difference between ICR- and INV-assessed median PFS, rounded to the nearest whole month. For comparative trials with more than two treatment groups, we calculated the difference in median PFS between the two treatment groups with the greatest difference in ICR-assessed median PFS. For comparative trials with two or three treatment groups, we calculated the difference in median PFS between each group reported by INV and ICR.

We reviewed recent oncology trials to establish the extent of differences between INV- and ICR-assessed PFS. We performed a search for studies where PFS was the primary endpoint (58%, 15) or co-primary endpoint (31%, 8) it was usually assessed by INV; PFS assessed by INV was the primary endpoint in 52% (14) of the studies and had reported results. Of these, 48 specified that PFS was assessed by INV, 45 (98%) by ICR, and 60 by both. The search of PubMed identified 49 studies, 21 of which reported both INV and ICR PFS values. There was one duplicate between the two databases, so the analysis set comprised 26 studies.

The majority of the study reports were for solid tumors, breast cancer was the most common cancer type (50%), followed by colorectal cancer (8%), lung (8%), and ovarian cancer (8%). Four hematological malignancies were included across 5 studies. Two trials (37%) had 3 groups, 14% had 4 groups, and 41% had 5 or more groups. A total of 43 comparisons were possible. We calculated the differences in median PFS as an endpoint and had reported results. Of these, 48 specified that PFS was assessed by INV, 45 (98%) by ICR, and 60 by both. The search of PubMed identified 49 studies, 21 of which reported both INV and ICR PFS values. There was one duplicate between the two databases, so the analysis set comprised 26 studies.

The difference between median PFS assessed by INV and ICR was calculated for each study and used as the outcome measure. The median PFS was calculated using the Kaplan-Meier method, and the log-rank test was used to assess the significance of the difference. The median PFS for each study was calculated using the Kaplan-Meier method, and the log-rank test was used to assess the significance of the difference. The median PFS for each study was calculated using the Kaplan-Meier method, and the log-rank test was used to assess the significance of the difference. The median PFS for each study was calculated using the Kaplan-Meier method, and the log-rank test was used to assess the significance of the difference.

### Results

A total of 365 Phase II or III interventional studies on ClinicalTrials.gov specified PFS as an endpoint and had reported results. Of these, 48 specified that PFS was assessed by ICR, 45 (98%) by ICR, and 60 by both. The search of PubMed identified 49 studies, 21 of which reported both INV and ICR PFS values. There was one duplicate between the two databases, so the analysis set comprised 26 studies.

The majority of the study reports were for solid tumors, breast cancer was the most common cancer type (35%), followed by colorectal cancer (9%), lung (8%), and ovarian cancer (8%). Four hematological malignancies were included across 5 studies. Two trials (37%) had 3 groups, 14% had 4 groups, and 41% had 5 or more groups. A total of 43 comparisons were possible. We calculated the differences in median PFS as an endpoint and had reported results. Of these, 48 specified that PFS was assessed by INV, 45 (98%) by ICR, and 60 by both. The search of PubMed identified 49 studies, 21 of which reported both INV and ICR PFS values. There was one duplicate between the two databases, so the analysis set comprised 26 studies.

The majority of the study reports were for solid tumors, breast cancer was the most common cancer type (35%), followed by colorectal cancer (9%), lung (8%), and ovarian cancer (8%). Four hematological malignancies were included across 5 studies. Two trials (37%) had 3 groups, 14% had 4 groups, and 41% had 5 or more groups. A total of 43 comparisons were possible. We calculated the differences in median PFS as an endpoint and had reported results. Of these, 48 specified that PFS was assessed by INV, 45 (98%) by ICR, and 60 by both. The search of PubMed identified 49 studies, 21 of which reported both INV and ICR PFS values. There was one duplicate between the two databases, so the analysis set comprised 26 studies.

The majority of the study reports were for solid tumors, breast cancer was the most common cancer type (35%), followed by colorectal cancer (9%), lung (8%), and ovarian cancer (8%). Four hematological malignancies were included across 5 studies. Two trials (37%) had 3 groups, 14% had 4 groups, and 41% had 5 or more groups. A total of 43 comparisons were possible. We calculated the differences in median PFS as an endpoint and had reported results. Of these, 48 specified that PFS was assessed by INV, 45 (98%) by ICR, and 60 by both. The search of PubMed identified 49 studies, 21 of which reported both INV and ICR PFS values. There was one duplicate between the two databases, so the analysis set comprised 26 studies.

The majority of the study reports were for solid tumors, breast cancer was the most common cancer type (35%), followed by colorectal cancer (9%), lung (8%), and ovarian cancer (8%). Four hematological malignancies were included across 5 studies. Two trials (37%) had 3 groups, 14% had 4 groups, and 41% had 5 or more groups. A total of 43 comparisons were possible. We calculated the differences in median PFS as an endpoint and had reported results. Of these, 48 specified that PFS was assessed by INV, 45 (98%) by ICR, and 60 by both. The search of PubMed identified 49 studies, 21 of which reported both INV and ICR PFS values. There was one duplicate between the two databases, so the analysis set comprised 26 studies.

The majority of the study reports were for solid tumors, breast cancer was the most common cancer type (35%), followed by colorectal cancer (9%), lung (8%), and ovarian cancer (8%). Four hematological malignancies were included across 5 studies. Two trials (37%) had 3 groups, 14% had 4 groups, and 41% had 5 or more groups. A total of 43 comparisons were possible. We calculated the differences in median PFS as an endpoint and had reported results. Of these, 48 specified that PFS was assessed by INV, 45 (98%) by ICR, and 60 by both. The search of PubMed identified 49 studies, 21 of which reported both INV and ICR PFS values. There was one duplicate between the two databases, so the analysis set comprised 26 studies.