Defining *EGFR* amplification status for clinical trial inclusion

Pim J. French, Marica Eoli, Juan Manuel Sepulveda, Iris de Heer, Johan M. Kros, Annemieke Walenkamp, Jean-Sebastien Frenel, Enrico Franceschi, Paul M. Clement, Michael Weller, Peter Ansell, Jim Looman, Earle Bain, Marie Morfouace, Thierry Gorlia, and Martin van den Bent

Departments of Neurology (P.J.F., I.D.H., M.V.D.B.) and Pathology (J.M.K.), Erasmus Medical Center Cancer Institute, Rotterdam, Netherlands; Carlo Besta Neurological Institute, Milan, Italy (M.E.); University Hospital, Madrid, Spain (J.M.S.); University Medical Center Groningen, Groningen, Netherlands (A.W.); Western Cancer Institute, René Gauducheau Center, Saint-Herblain, France (J.S.F.); Azienda USL/IRCCS Institute of Neurological Sciences, Bologna, Italy (E.F.); UZ Leuven, Leuven, Belgium (P.M.C.); Department of Neurology, University Hospital and University of Zurich, Zurich, Switzerland (M.W.); AbbVie, North Chicago, Illinois, USA (P.A., J.L., E.B.); European Organisation for Research and Treatment of Cancer (EORTC) Headquarters, Brussels, Belgium (M.M., T.G.)

Corresponding Author: Pim J. French, Dept Neurology, Erasmus MC, Rotterdam, The Netherlands (p.french@erasusmc.nl).

**Abstract**

**Background.** Precision medicine trials targeting the epidermal growth factor receptor (*EGFR*) in glioblastoma patients require selection for *EGFR*-amplified tumors. However, there is currently no gold standard in determining the amplification status of *EGFR* or variant III (*EGFRvIII*) expression. Here, we aimed to determine which technique and which cutoffs are suitable to determine *EGFR* amplification status.

**Methods.** We compared fluorescence in-situ hybridization (FISH) and real-time quantitative (RT-qPCR) data from patients screened for trial inclusion into the Intellance 2 clinical trial, with data from a panel-based next generation sequencing (NGS) platform (both DNA and RNA).

**Results.** By using data from >1000 samples, we show that at least 50% of *EGFR* amplified nuclei should be present to define *EGFR* gene amplification by FISH. Gene amplification (as determined by FISH) correlates with *EGFR* expression levels (as determined by RT-qPCR) with receiver operating characteristics analysis showing an area under the curve of up to 0.902. *EGFR* expression as assessed by RT-qPCR therefore may function as a surrogate marker for *EGFR* amplification. Our NGS data show that *EGFR* copy numbers can strongly vary between tumors, with levels ranging from 2 to more than 100 copies per cell. Levels exceeding 5 gene copies can be used to define *EGFR*-amplification by NGS; below this level, FISH detects very few (if any) *EGFR* amplified nuclei and none of the samples express *EGFRvIII*.

**Conclusion.** Our data from central laboratories and diagnostic sequencing facilities, using material from patients eligible for clinical trial inclusion, help define the optimal cutoff for various techniques to determine *EGFR* amplification for diagnostic purposes.

**Key Points**

1. We show which cutoffs define *EGFR* amplification for diagnostic purposes.
2. Diagnostic *EGFR* amplification cutoffs are defined for FISH, RT-qPCR, and NGS.
3. *EGFR* gene expression can serve as a surrogate marker for *EGFR* amplification.

In glioblastomas, the most common and aggressive type of primary brain tumor, the gene encoding epidermal growth factor receptor (*EGFR*) is mutated in approximately half of all tumors (Lassman et al, submitted). Initially, the *EGFR* locus is amplified to high copy number levels, sometimes exceeding 100 copies per tumor cell. Additional mutations
Importance of the Study

Precision medicine trials targeting EGFR in glioblastoma patients require selection for EGFR-amplified tumors. However, there is currently no standard to determine the amplification status of EGFR or EGFRvIII expression. In this study, we used molecular data derived from central laboratory testing and diagnostic-grade sequencing facilities, using a clinically relevant patient population (ie, those eligible for clinical trial inclusion) to help determine which technique and which cutoffs are suitable to determine EGFR amplification status for diagnostic purposes. Our results are of high relevance for the selection of patients into precision medicine trials.

Materials and Methods

Patients

Recurrent glioblastoma patients were considered eligible for the Intellance 2 trial (clinical trial identifier NCT02343406) if they had been diagnosed with a histologically confirmed, EGFR-amplified glioblastoma at first occurrence. For the screening assay, in the majority of cases (86%), material from first surgery was used. This is possible since EGFR amplification is a temporally stable genetic event in the large majority of patients. A total of 1084 samples were screened from which 260 patients were randomized in the Intellance 2 trial to receive either (i) TMZ or, if progressing within 16 weeks of the start of the last maintenance TMZ cycle, CCNU (n = 26 and n = 60, respectively; total n = 86); (ii) depatux-M alone (n = 86); or (iii) TMZ in combination with depatux-M (n = 88). A Consolidated Standards of Reporting Trials (CONSORT) diagram is shown in Figure 1. Eligibility criteria for molecular analysis were histologically confirmed de novo glioblastoma (primary) with unequivocal first progression after RT concurrent/adjacent chemotherapy at least 3 months post radiotherapy, age ≥18 years, and Karnofsky performance status 60–100. All centers had to obtain approval of the study design from their local ethical boards before study activation according to national and institutional regulations. All patients gave written informed consent for trial participation and molecular testing.

Fluorescence In-Situ Hybridization

FISH was performed by one of 3 laboratories (Histogenex, Antwerp, Belgium, n = 801; Mosaic, Lake Forest, California, n = 100; and Peter MacCallum Cancer Institute, Melbourne, Australia, n = 233) on glioblastomas using the Vysis EGFR CDx Assay (Abbott Molecular; not on market) as described (Lassman et al, submitted). Reproducibility between labs for FISH and RT-qPCR was determined using a proficiency tissue panel that was run on all sites upon startup. The proficiency of laboratory personnel to perform the Abbott RealTime EGFR assay was evaluated during RealTime EGFR assay training. The training/proficiency runs consist of a panel (glioblastoma formalin-fixed paraffin embedded [FFPE] tissue sections), EGFR positive and negative controls processed through all steps of the EGFR assay procedure (RNA isolation, amplification/detection reaction setup, and operation of the m2000 RealTime System). The trainees were evaluated based on performance of
One hundred percent concordance with expected results was one of the proficiency criteria. Data from all sites was evaluated by Abbott Molecular, and results were within expected range. For FISH, 2 DNA probes labeled with spectrally distinct fluorophores hybridizing to the 7p11.2-7p12 region or the chromosome enumeration probe (CEP) 7 hybridizing to the chromosome 7 centromere were used. Slides and probe mix were denatured at 73°C for 5 minutes and then hybridized at 37°C for 14–24 hours on a ThermoBrite system (Abbott Molecular). Sample pretreatment and post-hybridization washes were performed using the Vysis Universal FFPE Tissue Pretreatment and Wash Kit (Abbott Molecular; not commercially available). Slides were reviewed using fluorescence microscopy, and FISH signal counts (copy number) for individual probes were recorded for a total of 50 nuclei. A tumor was considered EGFR amplified when there was focal EGFR gene amplification defined as an EGFR/CEP 7 ratio greater than or equal to 2 in ≥15% recorded cells. Tumors with polysomy for chromosome 7 (excess copies of the entire chromosome) but without focal amplification of the EGFR gene were considered to be EGFR nonamplified.

**Real-Time Quantitative PCR**

RT-qPCR was performed as described (Lassman et al, submitted). Briefly, one ≥5 µM section containing a minimum of 50 mm² total tissue area from a FFPE tumor block was processed for RNA extraction using the Qiagen RNeasy FFPE Extraction Kit (Qiagen Sciences) according to the manufacturer’s instructions. FFPE sections were deparaffinized, and proteinase K was added. The nucleic acids were de-crosslinked from formalin and DNase treated to remove DNA content. Purified RNA was combined in a 96-well plate with mastermix containing primers and probes for amplification and detection of total EGFR and β-actin on the Abbott m2000 RealTime System (Abbott Molecular). β-actin served as an endogenous control and to provide relative quantitative values for total EGFR expression in the samples. The difference (ΔCt) between β-actin cycle threshold (Ct) and total EGFR Ct was calculated and reported.
Materials, either tissue sections or tissue blocks, were centrally collected at the Erasmus Medical Center in Rotterdam. Material of 226 patients included in the Intellance 2 trial was collected. Evaluation of the area with highest tumor content was done by the pathologist (J.M.K.) on a hematoxylin and eosin stained section. Ten to twenty 5-μm sections were then sent to Almac Diagnostics (Craigavon, UK) for macrodissection and subsequent DNA and RNA extraction using the Allprep DNA/RNA FFPE kit (Qiagen). Sequencing was done using the Trusight Tumor 170 panel (Illumina), which uses a combination of DNA and RNA sequencing to interrogate single nucleotide variations (SNVs) in ~150 genes, amplification of 59 genes, and fusion and splice variant expression in 55 genes. Sufficient quality was obtained in 216 samples for DNA sequencing and 215 for RNA sequencing. SNV, copy number, fusion gene, and splice variant expression calling was done on the Illumina Basespace sequence hub using the TruSight Tumor 170 App.

For each sample of the 1094 patients, one FISH assessment and one RT-qPCR value for EGFR, actin, and EGFRvIII was available. For 226 of these, the additional NGS data are available with one EGFR amplification level value (derived from a combination of probes on the EGFR locus, determined by the standard TruSight170 pipelines on the Illumina Basespace sequencing hub). EGFR amplification status by FISH was initially dichotomized as per trial inclusion (ie, using a cutoff of 15% amplified nuclei per tumor unless otherwise stated). EGFR amplification status as determined by RT-qPCR was dichotomized based on the bimodal distribution of the frequency histogram using ΔCt values stated in the analysis. Mean and cutoff values of bimodal distributions were determined based on a finite mixture model using the Expectation-Maximization algorithm and calculated based on the “cutoff” R package. This package was also identified the cutoff point for FISH to determine EGFR amplification. The ΔCt value with highest AUC (~3.56) and with a type I error rate of 0.05 (~2.48) was used to show concordance between FISH and RT-qPCR. Comparison between frequencies was done using the chi-square test. Kappa scores were calculated using the interrater reliability and agreement “irr” package in R. Where applicable, values listed with 95% confidence intervals or ±SD.

Results

For inclusion in the Intellance 2 trial, 1094 recurrent glioblastoma patients were screened for EGFR amplification by FISH in a central laboratory. Of these, 1072 samples yielded informative data to determine EGFR amplification (a CONSORT diagram is shown in Figure 1). Fifty nuclei were counted per sample from which the percentage of EGFR-amplified cells was calculated. For each cell, EGFR amplification was defined when the ratio of EGFR/centromere Chr 7 was >2. This would amount to a copy number of EGFR >6 per cell as most glioblastomas have gain of chromosome 7. Interestingly, and despite the fact that glioblastomas also harbor non-neoplastic stromal cells, the majority of samples either contain almost no EGFR amplified cells (<15% amplified nuclei) or consist entirely of EGFR amplified cells (>90% amplified nuclei; Figure 2).

![Fig. 2](https://academic.oup.com/neuro-oncology/article-abstract/21/10/1263/5498417)
Relatively few samples \((n = 122, 11.6\%)\) had intermediate numbers of EGFR-amplified nuclei. The FISH results therefore indicate that glioblastomas generally are not composed of mixtures of EGFR-amplified and non-amplified cells.

RT-qPCR was successfully performed on 1035 tumors. The histogram of \(\Delta \text{Ct}\) values, plotted in Figure 2B, shows a bimodal distribution with two peaks at \(-6.1 \pm 1.9\) and \(-1.4 \pm 1.4 \Delta \text{Ct}\). Ranked \(\Delta \text{Ct}\) values, plotted in Figure 3A, show a wide range of EGFR expression between different tumors. EGFR expression was correlated to amplification: the samples with high expression of EGFR almost always contained a high number of EGFR-amplified cells by FISH and, conversely, samples with low EGFR expression contained very few (if any) EGFR-amplified cells (Figure 3B). A receiver operating characteristic (ROC) plot shows the strength of gene expression in determining gene amplification (Figure 3C). When EGFR amplification was defined as a tumor having greater than \(-3.56 \Delta \text{Ct}\) EGFR versus actin expression value (ie, the intersect of 2 curves modeling the bimodal distribution), the area under the curve was 0.902 (95% CI: 0.881–0.922), with an optimum (ie, the point with the best sum of sensitivity and specificity) at sensitivity of 85.6% and specificity of 91.2% (95% CI: 88.1%–93.3%). At a sensitivity of 90% the specificity was 75.9% (95% CI: 60.6–88.8%). When EGFR amplification was defined as the optimal for type I errors (ie, 0.05), the \(\Delta \text{Ct}\) EGFR versus actin expression value cutoff increased to \(-2.48\). Also at this cutoff, the area under the curve remains high, 0.883 (95% CI: 0.862–0.904), with a sensitivity of 90.1% and specificity of 83.4%. The RT-qPCR data therefore demonstrate that EGFR expression can function as a surrogate marker to define EGFR amplification. Of note, in these analyses, the optimal threshold to determine EGFR amplification by FISH was a percentage of EGFR amplified cells exceeding 77%; higher than the 15% used for clinical trial inclusion.

The concordance between dichotomized FISH and RT-qPCR data was high: 89% using the RT-qPCR cutoff of \(-3.56 \Delta \text{Ct}\) values and 86% using the cutoff optimized for type I errors of \(-2.48 \Delta \text{Ct}\) values (both using a FISH cutoff of 77% amplified cells; Table 1). Of course, any increase in stringency to define EGFR amplification (by increasing the percentage of FISH positive cells) would reduce the number of EGFR-amplified tumors. Using the criteria for clinical trial inclusion (ie, >15% amplified cells), 565 of the 1035 samples tested (54.6%) were defined as being EGFR amplified. When increasing the stringency to define EGFR amplification to >77% amplified cells by FISH, this number would decrease to 460 (44.4%).

EGFRvIII expression was also determined by RT-qPCR and yielded informative data in 1049 samples. Almost all the 263 samples expressing EGFRvIII were also EGFR amplified as defined by FISH (257/263, 97.7%; \(P < 0.001\); Supplementary Figure 1A), and EGFRvIII expressing tumors were present mainly in samples with high levels of EGFR expression (Figure 4, Supplementary Figure 1B),

---

### Table 1  Concordance between FISH and RT-qPCR

<table>
<thead>
<tr>
<th>RT-qPCR Status (\Delta \text{Ct})</th>
<th>FISH Status (77%)</th>
<th>Negative</th>
<th>Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-3.56)</td>
<td>RT-qPCR status</td>
<td>69</td>
<td>506</td>
</tr>
<tr>
<td>(-2.48)</td>
<td>Positive</td>
<td>539</td>
<td>107</td>
</tr>
</tbody>
</table>

---

**Fig 3** RT-qPCR can function as a surrogate to determine EGFR amplification status. (A) Ranked \(\Delta \text{Ct}\) values from RT-qPCR data show a wide range of EGFR expression between samples. Samples with high EGFR expression often had EGFR amplification. (B) RT-qPCR plotted against the FISH data (% amplified cells). ROC curves show that EGFR expression can predict EGFR gene amplification by FISH. ROC curves were plotted for 4 different cutoffs (% of EGFR-amplified cells) to determine EGFR amplification by FISH.
confirming that EGFRvIII is expressed almost exclusively in samples with EGFR amplification.

**EGFR Amplification by NGS**

A total of 260 of the 1094 centrally screened patients were included in the Intellance 2 randomized phase II clinical trial. Patients were selected based on the presence of EGFR amplification by FISH only, with amplification defined as >15% of nuclei containing an EGFR/centromere Chr 7 ratio >2. Targeted NGS using the Illumina Trusight 170 platform was successfully performed (DNA) on 226 of these 260 samples. Ranked copy number estimates from the sequencing data are plotted in Figure 5A and confirms that most samples indeed contain high copy gene amplification. However, of the 226 samples sequenced, 30 samples (13.9%) were estimated to have fewer than 4 EGFR copy numbers. Most of these 30 low copy number samples also contained few EGFR-amplified cells as determined by FISH, and had low EGFR gene expression levels (Figure 5B). The observation that samples with a low percentage of amplified nuclei often show no copy number alterations by NGS suggests that FISH should use a more stringent cutoff to molecularly determine EGFR amplification than the threshold used for this trial. Indeed, when EGFR amplification was redefined as having >50% amplified nuclei, the total number of EGFR amplified samples drops marginally (n = 199/226), but the proportion of samples containing fewer than 4 EGFR copy numbers by NGS was reduced 29 to 7. It should be noted, however, that cutoffs used to molecularly define EGFR amplification may not reflect clinical efficacy of targeted agents; precision medicine trials may use a different cutoff.

It is possible that these samples are actually EGFR amplified but had a low tumor cell content. However, all sections were marked for areas of high tumor content (>70% tumor cells) by a dedicated neuropathologist. Moreover, the ranked copy number variation data from NGS show a “shoulder” in EGFR copy number variation between <5 and >10 gene copies (Figure 5A). This shoulder can be explained by the fact that EGFR is present as double minutes, which facilitates rapid high copy number gains, and thus argues against incorrect estimation of tumor percentage (when a smoother line may be expected). Finally, none of the 29 samples express EGFRvIII when EGFR copy numbers are low (<4), whereas the percentage is markedly higher, 103 of 187 (55.1%), in samples with >4 copy

---

**Fig. 4** EGFRvIII expression is found in samples with high levels of EGFR expression. As can be seen, samples that express EGFRvIII (y-axis) are found in samples that have high levels of EGFR expression (x-axis). Dark-gray dots indicate those that have been classified as EGFRvIII positive tumors (ie, Ct values of EGFRvIII <30).
French et al. Defining EGFR amplification status
Neuro-Oncology

Defining EGFR amplification status

EGFR amplification precedes the rearrangement leading to EGFRvIII expression, and generally half of EGFR-amplified tumors express EGFRvIII. Therefore, the absence of EGFRvIII-expressing cells in samples with <5 copy numbers (and <50% FISH positive cells) also suggests that the majority of these samples are incorrectly classified, with the used FISH cutoff, as EGFR amplified. Our data therefore strongly suggest that the cutoff to determine EGFR gene amplification by FISH should be increased to a sample having >50% EGFR-amplified cells.

Basic patient demographics (age, sex, etc) were only collected for samples that were included in the Intellance 2 trial. Within this group, demographics between RT-qPCR-high and RT-qPCR-low cohorts were similar for patients for which NGS was attempted (Supplementary Table 1).

The Trusight 170 platform also includes targeted RNA sequencing of ~60 genes (including EGFR) and was successfully performed on 215 samples. Read depth of EGFR in this cohort was high; the median number of reads (combined reference and splice supporting reads) was 37,000 (interquartile range, 23,000–60,000). Comparison in EGFR expression between RT-qPCR and RNA-seq shows a high concordance between the 2 techniques, with most samples that have low expression on both platforms also containing relatively few EGFR amplified nuclei by FISH (Figure 6A). EGFRvIII expression was detected in 107 and 93 samples by RT-qPCR and RNA-seq respectively, with 91 samples identified by both platforms, unweighted kappa score 0.829. The observation that RT-qPCR identifies more EGFRvIII-expressing samples may suggest this technique has a higher sensitivity to detect these aberrant transcripts, even despite the high read depth (Figure 6B).

Discussion

In the era of targeted treatments for cancer patients, identifying the target with an appropriate biomarker assay has become an essential part of both clinical studies and standard of care once shown effective. The biomarker assay used should be well validated: robust, reproducible, feasible, and predictive for outcome. The present study reviews 3 different assays to identify EGFR-amplified tumors, comparing 2 assays in the screened population and another technique for the randomized patients. In addition, the presence or absence of the EGFRvIII mutation is studied. EGFRvIII is a genomic rearrangement that occurs after EGFR amplification.

FISH for EGFR amplification is usually seen as a gold standard for the diagnostics of EGFR amplification, but what cutoffs (the % of positive nuclei, and the optimal ratio of EGFR vs centromere 7) should be used is unclear. For this trial, the bar to call EGFR amplification was set low, in view of the observation of a response in a patient with a relatively low level of EGFR amplification. An advantage of
Defining EGFR amplification status

such a low bar in a targeted treatment trial is that it allows the assessment of efficacy of the investigational compound around the borders of the assay. By using data from >1000 samples, we provide evidence that the cutoff to determine EGFR gene amplification by FISH should be increased, preferably to >50% of EGFR-amplified nuclei. Although the number of samples in the current study was large, validation of this observation in an independent cohort would strengthen our conclusions. Such cohorts are, however, at present unavailable.

Although this FISH cutoff may better determine EGFR amplification for diagnostic purposes, response prediction in precision medicine trials (ie, the predictive power of EGFR amplification) may use different cutoff points. For example, when EGFR-amplified cells exert a dominant effect on the non-amplified neoplastic cells (whereby the non-amplified neoplastic cells depend on the EGFR-amplified cells), targeting the minority of EGFR-amplified cells may already provide clinical benefit. The observation of patients with relatively low levels of EGFR amplification respond to the combination of deputux-M + TMZ supports this notion.

Our data also demonstrate that EGFR expression can serve as a surrogate marker to determine EGFR gene amplification. However, where EGFR-FISH has a bimodal distribution, EGFR expression is much more a continuum. This continuum makes it more difficult to define the optimal cutoff for EGFR amplification. Of note, the bimodal distribution observed by FISH demonstrates that most tumors either have no amplified cells or consist entirely of amplified cells. This indicates that the frequency of mosaic amplification of various oncogenes as previously reported is therefore likely relatively low.27

In summary, the data described in the current study, obtained from central laboratories and diagnostic sequencing facilities and using material from patients eligible for clinical trial inclusion, help define the optimal cutoff for various techniques to determine EGFR amplification for selection into precision medicine trials. In the end, the most optimal cutoff should be established based on evidence of clinical activity in the positive versus the negative biomarker population. That analysis is currently ongoing. Regardless, although our study shows that these tests correlate, there is a clear gray zone in which samples test positive in one assay but not in another assay.

Supplementary Material
Supplementary data are available at Neuro-Oncology online.

Keywords
amplification | biomarker | EGFR | EGFRvIII | FISH | glioblastoma | screening

---

**Fig. 6** RT-qPCR and NGS show high correlation in determining levels of EGFR and presence of EGFRvIII. (A) EGFR expression by RT-qPCR is highly correlated with expression derived from NGS data. (B) Samples in which EGFRvIII expression is identified by RT-qPCR (ie, ΔCt < 30) often also show detectable levels of EGFRvIII by panel-based RNA-seq.
Funding

This research was sponsored and funded by AbbVie Inc.

Conflict of interest statement. PJJ received research funding by AbbVie, MJvdB consultancy for AbbVie, Cellgene, Boehringer, BMS, AGIOS, PA, JL and EB are employees of AbbVie and may own stock. MW has received research grants from AbbVie, Adastra, Bayer, Merck, Sharp & Dohme (MSD), Dracen, Merck (EMD), Novocure, OGD2, Piqur and Roche, and honoraria for lectures or advisory board participation or consulting from AbbVie, Basilea, Bristol Meyer Squibb, Cellgene, Merck, Sharp & Dohme (MSD), Merck (EMD), Novocure, Orbis, Roche, Teva, and Tocagen. JMS received research funding from Pfizer and Catalysis, consulting or advisory board for Abbvie, Cellgene, and Pfizer, and travel expenses from Astellas and Abbvie. PC served on advisory boards for Abbvie.

Authorship statement. Conceptualization, PJF, MvdB. Methodology, PJF. Investigation, IDh, JMK. Writing, original draft, PJF. Writing, review, and editing, all authors. Funding acquisition, PJF. Resources, ME, J-MS, AW, J-SF, EF, PMC, MW, PA, JL, EB. Data curation: MM, TG. Supervision, PJF. Writing, review, and editing, all authors. Funding acquisition, PJF. Conceptualization, PJF, MvdB.

References


