SUMMARY OF SAFETY AND EFFECTIVENESS DATA (SSED)

I. GENERAL INFORMATION

Device Generic Name:	Real-time PCR test
Device Trade Name:	*therascreen*® FGFR RGQ RT-PCR Kit
Device Procode:	OWD
Applicant's Name and Address:	QIAGEN Manchester Ltd Skelton House, Lloyd Street North, Manchester, M15 6SH, United Kingdom
Date(s) of Panel Recommendation:	None
Premarket Approval Application (PMA) Number	P180043
Date of FDA Notice of Approval:	April 12, 2019

II. INDICATIONS FOR USE

The *therascreen* FGFR RGQ RT-PCR Kit is a real-time, reverse transcription PCR test for the qualitative detection of two point mutations in exon 7 [p.R248C (c.742C>T), p.S249C (c.746C>G)], two point mutations in exon 10 [p.G370C (c.1108G>T) and p.Y373C (c.1118A>G)] and two fusions (FGFR3-TACC3v1 and FGFR3-TACC3v3) in the fibroblast growth factor receptor 3 (FGFR3) gene in RNA samples derived from formalin-fixed paraffin-embedded (FFPE) urothelial tumor tissue. The test is indicated for use as an aid in identifying urothelial cancer (UC) patients who harbor these alterations and are therefore eligible for treatment with BALVERSA™ (erdafitinib).

Specimens are processed using the RNeasy DSP FFPE Kit for manual sample preparation followed by reverse transcription and then automated amplification and detection on the Rotor-Gene Q MDx (US) instrument.

III. CONTRAINDICATIONS

There are no known contraindications.

IV. WARNINGS AND PRECAUTIONS

The warnings and precautions can be found in the *therascreen*® FGFR RGQ RT-PCR Kit labeling.

V. DEVICE DESCRIPTION

The *therascreen*® FGFR RGQ RT-PCR Kit (FGFR Kit) is a two-step reverse transcription, real-time PCR to detect the nine alterations in the FGFR2 and FGFR3 genes detailed in Table 1. Each kit contains reagents to test up to 24 patient specimens, plus additional reagents to test control samples.

Table 1: FGFR Alterations in the FGFR Kit

	Gene	Amino acid variant	CDS Mutation	Cosmic ID	Exons
Point Mutations	FGFR3	p.R248C	c.742C>T	COSM714	7
	FGFR3	p.G370C	c.1108G>T	COSM716	10
	FGFR3	p.S249C	c.746C>G	COSM715	7
	FGFR3	p.Y373C	c.1118A>G	COSM718	10
Fusions	**Fusion ID**		**Genes involved**	**Genomic breakpoints**	**Exons**
	FGFR3-TACC3v1		FGFR3	chr4:1808661 C	17
			TACC3	G chr4:1741428	11
	FGFR3-TACC3v3		FGFR3	chr4:1808661 C	17
			TACC3	G chr4:1739324	10
	FGFR3-BAIAP2L1*		FGFR3	chr4:1808661 C	17
			BAIAP2L1	A chr7:97991744	2
	FGFR2-BICC1*		FGFR2	chr10:123243211 G	17
			BICC1	A chr10:60461834	3
	FGFR2-CASP7*		FGFR2	chr10:123243211 G	17
			CASP7	A chr10:115457252	2

* The test was designed to identify FGFR2 fusions (FGFR2-BICC1 and FGFR2-CASP7) and FGFR3 fusion FGFR3-BAIAP2L1 because these FGFR fusion patients were eligible for the trial, however, the Qiagen test is not clinically validated to detect these 3 fusions due to the lack of required clinical sample materials (see clinical section X for additional information).

The following components comprise the overall device:

- QIAGEN RNeasy® DSP FFPE Kit – for manual RNA extraction and purification
- QIAGEN *therascreen*® FGFR RGQ RT-PCR Kit (FGFR Kit) – for reverse transcription and real-time PCR
- QIAGEN Rotor-Gene® Q MDx Instrument (RGQ)
- Rotor-Gene AssayManager (RGAM) Software Version 2.1
- Rotor-Gene AssayManager Gamma MDx plug-in version 1.0
- *therascreen*® FGFR FFPE MDx Assay Profile v1.0.0

A. Specimen Preparation

Initial preparation of specimens for use with the FGFR Kit follows standard pathology procedures. Tumor tissue is typically fixed in 10% neutral buffered formalin and then embedded in paraffin to form FFPE blocks. The FFPE blocks must be sectioned using a microtome, Hematoxylin and Eosin (H&E) stained and reviewed by a qualified pathologist to assess tumor area, content and composition. The FFPE tissue input for the RNeasy® DSP FFPE Kit used with the FGFR Kit must be made equivalent to a 4-5 µm slide thickness with a total tumor area between 100 mm^2 and 500 mm^2, which can be created from multiple slides. Macrodissection is performed (if necessary) to attain a minimum of 80% tumor surface area. If tumor area is below 100 mm^2 and/or the tumor -surface area lower than 80%, additional sections must be used to reach the minimum sample requirements. A minimum of 10% viable tumor cells are required. Non-tumor areas (e.g., necrotic tissue) identified by the pathologist must be removed by macrodissection if it constitutes greater than 20% of the section. RNA is extracted from the FFPE specimens using the RNeasy® DSP FFPE Kit.

B. Qualification and Normalization of RNA Samples

The RNA input concentration in the reverse transcription reaction is fixed at 250 ng in 15 µl. To attain this concentration, after extraction RNA samples are normalized to a fixed concentration of 16.67 ng/µl by dilution with the "Water for Sample Dil." supplied in the FGFR Kit. If the RNA concentration is below 16.67 ng/µl, the sample must not be processed further. A fresh RNA extraction from a new FFPE specimen should be used for further analysis. The RNA normalization process is performed at 2°C to 8°C.

C. Reverse Transcription, PCR Amplification and Detection

Reverse Transcription

To carry out the test procedure, complementary DNA (cDNA) is first synthesized from sample RNA using Reverse Transcriptase. It is this cDNA which then acts as the initial template in PCR.

Reverse transcription is performed using a master mix prepared from Reverse Transcriptase enzyme solution, RT Buffer1, RT Buffer 2 and RT Primer Mix, which are all provided with the FGFR Kit. The reverse transcription reaction takes place in a suitable heating block, water bath or thermal cycler by incubating the reaction mix at 42°C and then inactivating the reaction at 95°C.

Reverse transcriptase is a multifunctional enzyme with three distinct enzymatic activities: an RNA-dependent DNA polymerase, a hybrid-dependent exoribonuclease (RNase H) and a DNA-dependent DNA polymerase. For reverse transcription in vitro the first two activities are utilized to produce single-stranded cDNA. Firstly, the RNA-dependent DNA-polymerase activity transcribes cDNA from an RNA template, forming a DNA:RNA hybrid. Subsequently, the RNase H exonuclease activity specifically

degrades only the RNA strand of these hybrids. Therefore, this activity affects RNA hybridized to cDNA, but has no effect on pure RNA. A separate RNA degradation step using RNase H enzyme is not necessary prior to real-time PCR.

Real Time PCR Amplification

Primers: For the selective detection of the nine FGFR2 and FGFR3 alterations, oligonucleotide primer pairs (forward and reverse) are designed to flank each of the specific alterations (either the point mutation or the exon-exon spanning region), generating amplicons shorter than 120 base pairs (bp). Each primer has on average 20 bp (maximum 24 bp, minimum 16 bp) and a GC content of approximately 50%.

Allele-specific TaqMan® Probes: TaqMan® technology - a detection system based on hydrolysis probes - is used to detect and measure the amplified cDNA product. The probes cover the mutation site to detect point mutations, or the exon-exon spanning region to detect the fusion alterations.

In the real-time PCR reaction, forward and reverse primers hybridize to a specific cDNA sequence, or target sequence, to amplify the PCR product. TaqMan® probes are also contained in the reaction mixes. Each probe has a target allele-specific design, meaning it hybridizes exclusively to a target sequence. When the wild-type allele or an alternative allele is present in the sequence, the probe does not hybridize. Each probe, which consists of an oligonucleotide labeled with a 5' reporter dye (e.g. FAM) and a downstream, 3' dye-free quencher (e.g. BHQ1), also hybridizes to the target sequence between the forward and reverse primers.

The hydrolysis probe method exploits the 5'→3' exonuclease activity of the (*Taq*) DNA polymerase. When the probes are intact, the proximity of the reporter dye to the quencher results in suppression of the reporter fluorescence primarily by energy transfer. During real-time PCR, both forward and reverse primers anneal to the target cDNA and, depending on the target sequence present, a probe will specifically bind the target sequence located between the primers. The 5'→3' exonuclease activity of the polymerase cleaves the probe between the reporter and the quencher only if the probe has specifically bound to its target sequence. The probe fragments are then displaced from the target, leading to an increase in detectable reporter fluorescence.

The 3' end of the probe is blocked to prevent extension of the probe during real-time PCR. This process occurs in every cycle and does not interfere with the exponential accumulation of product. The increase in fluorescence signal is detected only if the target sequence, which is complementary to the probe, is present. Thus, the increase in fluorescence is directly proportional to the target amplification during real-time PCR.

PCR: The FGFR Kit contains reagents that allow PCR amplification and qualitative detection of the alterations listed in Table 1. As a fixed RNA amount is required for the reverse transcription reaction (250 ng), an internal control (IC) assay functions to determine if an appropriate level of amplifiable cDNA is present in the sample. Each

reaction mix contains a set of IC primers and probes designed to target a conserved region of the b-actin (ACTB) gene, a constitutively and stably expressed housekeeping gene in the transitional epithelium.

The probes used in the alteration-specific reaction mixes are labeled with carboxyfluorescein (FAM), Cal Fluor® Red 610 and cyanine (Cy5.5) fluorescent reporter dyes, each with a distinct absorption and emission profile. The probe used in the IC reaction is labeled with hexachlorofluorescein (HEX).

FAM, HEX, Cal Fluor® Red 610 and Cy5.5 absorb and fluoresce at different wavelengths:

- FAM: a fluorophore that excites at a wavelength of 495nm and emits at a wavelength of 520nm. This fluoresces in the Green RGQ channel.
- HEX: a molecule that becomes excited at a wavelength of 535nm and fluoresces at a wavelength of 556nm. This fluoresces in the Yellow RGQ channel.
- Cal Fluor® Red 610: An amidite that excites at 590 nm and emits at 610 nm. This fluoresces in the Orange RGQ channel.
- Cy5.5: a fluorophore of the cyanine family that excites at 675 nm and emits at 694 nm. This fluoresces in the Crimson RGQ channel.

PCR amplification for each allele and the IC can therefore be detected simultaneously.

The Mut-1 and Mut-2 reaction mixes also contain wild-type blockers, which are short oligonucleotide sequences with modifications in the 3' end that prevent sequence extension by *Taq* polymerase.

PCR Cycle Threshold Determination

The number of PCR cycles necessary to detect a fluorescent signal above a pre-determined threshold is called the cycle threshold (C_T) and is inversely proportional to the amount of target present at the beginning of the reaction, allowing a sensitivity limit to be set for the test.

The FGFR alteration status of a sample is determined based on the obtained C_T values for the analyzed assays. Samples are classed as "FGFR Alteration Detected" if they yield a C_T at or below the cut-off C_T value for a particular assay. Above this value, the sample may either contain less than the copy number able to be detected by the FGFR Kit (beyond the limit of detection of the Kit), or the sample is negative for an FGFR alteration, both of which are reported as "No Alteration Detected". Results for each sample are determined based on C_T values. The RNA quality is checked through the IC C_T values, which should be within the working range defined in the sample acceptance criteria.

D. Test Controls

The FGFR Kit contains three controls: a Positive Control (PC), an Internal Control (IC), and a No Template Control (NTC) which have been designed to detect fault conditions.

Internal Control (IC): Each multiplex reaction mix contains a set of primers and probes designed to target a conserved region of the ACTB gene, a constitutively and stably expressed housekeeping gene in the transitional epithelium. As a fixed RNA amount is required for the reverse transcription reaction (250 ng), the housekeeping gene functions as an IC to determine if an appropriate level of amplifiable cDNA is present in the sample after RNA extraction and reverse transcription.

No Template Control (NTC): The NTC is nuclease-free water and its use is common practice in laboratories using real-time PCR. An NTC is included on each run to assess the presence of contaminants and to determine run validity. The NTC is included in the reverse transcription step and carried through to the real-time PCR step.

Positive Control (PC): The FGFR Kit includes a PC to confirm the proper functioning of the reagents in the reverse transcription and real-time PCR steps. The PC is an in vitro transcript (IVT) that carries all of the alterations detected by the FGFR Kit, including the IC. Using an IVT template enables the PC to control for the function of both the reverse transcription and real-time PCR steps. The matrix for the PC is a buffer containing Tris-EDTA and Poly-A carrier RNA. This is a standard IVT storage buffer designed to maintain stability of the control material. The Poly-A carrier RNA acts to stop the RNA IVT from binding to the plastic tube. The IVT is diluted in PC Diluent, which is a solution of human universal RNA, prior to addition to the reverse transcription reaction mix. The resulting cDNA is then carried forward to the PCR reaction.

E. Instrument and Software

The real-time PCR step of the FGFR Kit is designed to be used with the RGQ instrument which is a real-time PCR analyzer designed for rapid thermal cycling and real-time detection of PCR assays.

The RGAM software, associated plug-in and *therascreen*® FGFR FFPE MDx Assay Profile control and monitor real-time PCR reactions and allow the determination of the diagnostic status based upon the real-time PCR results.
The RGQ incorporates a centrifugal rotary design for thermal cycling where each tube spins in a chamber of circulating air, keeping all samples at a uniform temperature. Samples are heated and cooled in a low-mass-air oven per a software determined cycle that initiates the different phases of the real-time PCR cycling profile. In the RGQ, fluorophores are excited from the bottom of the sample chamber by a light-emitting diode. Energy is transmitted through the thin wall at the bottom of each real-time PCR tube. Emitted fluorescence passes through the emission filters on the side of the chamber and is detected by a photomultiplier tube. Detection is performed as each tube aligns

with the detection optics; tubes spin past the excitation / emission optics every 150 milliseconds. The fluorescence signals indicate the progress of the real-time PCR reactions. The RGQ has five optical channels (five excitation sources and five detection filters). For the FGFR Kit, four of these channels are used: green, yellow, orange and crimson.

The RGAM software is a core software used with the FGFR Kit which provides general functionality including real-time PCR run setup, cycler control and management of experiment data, results, assay profiles and system configuration. The core application is complemented by plug-ins which can perform assay specific analysis, result visualization and reporting. The assay profile provides configuration information for the RGAM core and plug-in to enable assay-specific functionality. The RGAM software therefore controls the basic operation of the RGQ instrument, independent of the assays being used with it, while the plug-ins with the assay profiles provide data analysis and result reporting.

The RGAM Gamma MDx plug-in extends the functionality of RGAM by providing C_T value calculation, data analysis and run analysis normalization features. The plug-in is not specific to the FGFR Kit; instead it is designed to provide general functionality to assays or assay groups.

The RGAM software supplied for use with the FGFR Kit contains additional analysis features to those contained in the standard RGQ software. In order to use the FGFR-specific features, a compatible plug-in must be installed on the RGQ together with the *therascreen*® FGFR FFPE MDx Assay Profile.

The RGAM software and associated plug-in and assay profile ensure that a user interface with restricted user options is displayed to the user and contain all the information required for automatic real-time PCR analysis. In addition, the software suite allows printing of test reports and creates result files in the software's file system.

Assay specific functionality is implemented by the *therascreen*® FGFR FFPE MDx Assay Profile, for example cycling conditions, thresholds and analysis cut-offs, and control ranges.

F. Interpretation of Results

Based on the C_T values obtained for controls and samples, the RGAM software determines if the run controls are valid, if individual samples give valid results and reports the FGFR alteration status of valid sample tests only.

For FGFR Kit runs to be accepted as valid, the RGQ software requires run data for PC and NTC must meet criteria specified in the *therascreen*® FGFR FFPE MDx Assay Profile. All run validity criteria must be met before a run is considered valid and individual sample data are analyzed. The run validity criteria for a RGQ run to be considered valid are detailed in Table 2 for PC and Table 3 for NTC.

Table 2: Run Validity Criteria for PC

Reaction Mix	Alteration	C_T values [(a)]	
		Lower Limit	Upper Limit
Mut-1	p.R248C (c.742C>T)	22.98	28.58
	p.G370C (c.1108G>T)	24.41	30.07
	IC (ACTB)	16.33	21.94
Mut-2	p.S249C (c.746C>G)	22.67	30.92
	p.Y373C (c.1118A>G)	25.55	34.45
	IC (ACTB)	16.24	21.84
Fus-1	FGFR3-TACC3v3	23.39	29.2
	FGFR3-BAIAP2L1	22.66	28.26
	FGFR2-CASP7	23.23	28.85
	IC (ACTB)	16.01	21.62
Fus-2	FGFR3-TACC3v1	22.79	29.85
	FGFR2-BICC1	24.55	30.15
	IC (ACTB)	15.84	21.44

(a) *Upper and lower C_T values are inclusive (i.e. include the values shown)*

Table 3: Run Validity Criteria for NTC

Reaction Mix	Alteration	All levels of testing
Mut-1	p.R248C (c.742C>T)	$C_T > 43.00$
	p.G370C (c.1108G>T)	$C_T > 43.00$
	IC (ACTB)	$C_T > 43.00$
Mut-2	p.S249C (c.746C>G)	$C_T > 39.09$
	p.Y373C (c.1118A>G)	$C_T > 43.00$
	IC (ACTB)	$C_T > 43.00$
Fus-1	FGFR3-TACC3v3	$C_T > 43.00$
	FGFR3-BAIAP2L1	$C_T > 43.00$
	FGFR2-CASP7	$C_T > 43.00$
	IC (ACTB)	$C_T > 43.00$
Fus-2	FGFR3-TACC3v1	$C_T > 43.00$
	FGFR2-BICC1	$C_T > 43.00$
	IC (ACTB)	$C_T > 43.00$

If a run fails any of the run validity criteria, the RGAM software displays the corresponding validity rule related to the failed control but does not provide the test results for samples on the RGAM report. If all run validity criteria are correct, the RGAM generates a report that confirms the respective controls validity and then displays the samples results (if sample release criteria are met). If the run passes the run validity criteria, the RGAM software then uses the C_T values compared to the assay cut-offs to determine the sample alteration status.

As a fixed RNA amount is required for the reverse transcription reaction (250 ng in 15 µl) and 5 µl cDNA is used for each real-time PCR multiplex reaction mix, the housekeeping gene functions as an IC to determine if an appropriate level of amplifiable cDNA is present in the sample. The FGFR Kit has been demonstrated to work within a

specific working range (i.e. upper and lower IC C_T values). Samples are deemed valid if the IC C_T values are within the working range and samples that do not give C_T values within this range are invalidated by the RGAM software.

The FGFR Kit IC working range was determined in the FGFR Kit through the analysis of the range of C_T values obtained with this assay when testing 250 ng of RNA derived from clinical FFPE UC specimens. The working range specifications are provided in Table 4.

Table 4: Internal Control Working Range

Control C_T value	Interpretation	Action
> 25.58	Quantity of amplifiable cDNA is not sufficient for FGFR alteration analysis	Additional samples should be extracted and tested
< 17.47	Quantity of amplifiable cDNA is too high for FGFR alteration analysis	Additional samples should be extracted and tested
$17.47 \leq$ Control $C_T \leq 25.58$	Quantity of amplifiable cDNA is suitable for FGFR alteration analysis	No action required, sample is suitable

Samples which give an IC C_T value within the IC working range, i.e. produce a C_T value of 17.47 – 25.58, are deemed valid. Samples with an IC C_T value outside this range are classified as invalid, due to sample excess or degradation.

If the RGAM software fails to detect a signal within the working range in the IC assay for a specific reaction mix, the sample is reported as invalid and no FGFR result is reported.
On the RGAM report each sample is assigned with a status as follows:

INVALID:
If one of the run control criteria failed
or if any IC C_T value for a specific reaction mix is outside the specific working range.

FGFR Alteration Detected:
If all run and sample control criteria are met for a particular sample
and the C_T value for at least one of the alterations is below the assay cut-off

No Alteration Detected:
If all run and sample control criteria are met for a particular sample
and C_T value for all the alterations is above the assay's cut-offs, or no amplification is detected.

VI. ALTERNATIVE PRACTICES AND PROCEDURES

There are no other FDA-cleared or approved alternatives for FGFR alteration testing of FFPE UC tissue for the selection of patients who are eligible for treatment with BALVERSA™ (erdafitinib).

VII. MARKETING HISTORY

The Qiagen *therascreen* ® FGFR RGQ RT-PCR Kit has not been marketed in the United States or any foreign country.

VIII. POTENTIAL ADVERSE EFFECTS OF THE DEVICE ON HEALTH

Failure of the device to perform as expected or failure to correctly interpret test results may lead to incorrect FGFR test results and subsequently improper patient management decisions in Urothelial Cancer (UC) treatment.

For the specific adverse events related to BALVERSA™ (erdafitinib) that occurred in the clinical studies, please see Section X below. The most frequently reported treatment emergent adverse events (TEAEs) in the selected regimen (8 mg daily) of the clinical trial were hyperphosphatemia, stomatitis, diarrhea, dry mouth, decreased appetite, dry skin, fatigue, and dysgeusia. Adverse events of clinical concern are those expected of agents that inhibit FGFRs, mainly eye, nail, and skin disorders as well as hyperphosphatemia.

IX. SUMMARY OF NONCLINICAL STUDIES

A. Laboratory Studies

Procured FFPE tissue blocks and sections obtained from patients with advanced UC, and contrived samples were used in the analytical studies. Clinical samples were macrodissected when the tumor surface area in the sample was less than 80% or diluted according to the instructions to users. The similarity between clinical FFPE specimens and contrived samples was demonstrated for four representative alterations of the *therascreen* ® FGFR RGQ RT-PCR Kit (hereafter FGFR Kit) (p.R248C (c.742C>T), p.Y373C (c.1118A>G), FGFR3-TACC3v1 and FGFR3-TACC3v3) by comparing positivity rates and probit models for the two sample types.

1. Correlation to Comparator Method/Accuracy

To demonstrate the accuracy of the FGFR Kit relative to an analytically validated high-throughput orthogonal comparator method, an accuracy study was conducted with specimens from the 42756493BLC2001 clinical trial, supplemented with procured specimens from the same intended use population. The FGFR Kit and comparator testing for FGFR alterations was performed on the same RNA samples extracted from 307 FFPE specimens (including 271 clinical trial specimens with sufficient RNA from the Bridging Study and 36 procured specimens).

To evaluate concordance, the positive percentage agreement (PPA), negative percentage agreement (NPA), and overall percent agreement (OPA), along with the respective two-sided 95% confidence intervals (CIs) were calculated. The measures of agreement were calculated using the comparator as the reference method. As such, PPA was calculated as the proportion of FGFR Kit positive (+) given that the sample was comparator+; NPA was calculated as the proportion of FGFR Kit negative (-) given that the sample was comparator-, and OPA was calculated as the proportion of agreement between the FGFR Kit and comparator among all samples tested.

Three hundred and six (306) subjects provided a valid test result. The two-by-two frequency table (excluding invalid results) of FGFR Kit overall alteration status against comparator method alteration status is presented in Table 5.

Table 5: Two-by-Two Frequency Table (Excluding Invalid Results) of the FGFR Kit vs. Comparator Call

		Comparator Method		
FGFR Kit	**Frequency**	**Negative**	**Positive**	**Total (Percent)**
	Negative	197	1	198 (64.71)
	Positive	5	103	108 (35.29)
	Total (Percent)	202 (66.01)	104 (33.99)	306 (100.00)

The estimated PPA, NPA and OPA between the FGFR Kit and comparator, with comparator as the reference method, were 99.04%, 97.52% and 98.04%, respectively (Table 6).

Table 6: The PPA, NPA and OPA with comparator result as the reference method along with the two-sided 95% confidence intervals

Measure of Agreement	Percent agreement % (N)	Clopper-Pearson (Exact) Binomial Lower, Upper Two-sided 95% CI
PPA	99.04% (103/104)	94.76, 99.98
NPA	97.52% (197/202)	94.32, 99.19
OPA	98.04% (300/306)	95.78, 99.28

For the six overall FGFR alteration status discordant results, one sample yielded a "No Alteration Detected" result by the FGFR Kit and gave "FGFR Alteration Detected" result(s) by comparator method, while five samples gave "FGFR Alteration Detected" result(s) by the FGFR Kit and gave "No Alteration Detected" results by comparator method.

Agreements with comparator method by each mutation are listed in Table 7 below. A total of 104 comparator positive subjects harbored 113 genetic alterations.

Table 7: The PPA with a comparator method along with the two-sided 95% confidence intervals by alterations

Alteration	Percentage agreement % (N correct/N Total)	Clopper-Pearson (Exact) Binomial Lower, Upper Two-sided 95% CI
p.R248C (c.742C>T)	93.33% (14/15)	68.05, 99.83
p.S249C (c.746C>G)	100.00% (56/56)	93.62, 100.00
p.G370C (c.1108G>T)	100.00% (2/2)	15.81, 100.00
p.Y373C (c.1118A>G)	100.00% (18/18)	81.47, 100.00
FGFR3-TACC3v1	100.00% (16/16)	79.41, 100.00
FGFR3-TACC3v3	100.00% (5/5)	47.82, 100.00
FGFR3-BAIAP2L1	100.00% (1/1)	2.50, 100.00

2. Analytical Sensitivity

a) Limit of Blank (LoB)

To ensure that a sample with wild-type RNA does not generate an analytical signal that might indicate a low concentration of alteration, RNA from 60 individual FGFR wild-type FFPE UC specimens was tested at the fixed input concentration specified for the FGFR Kit. A single replicate of each RNA sample was tested with three FGFR Kit lots using three RGQ instruments over three days, generating a total of 180 data points for each assay. No results were reported as invalid and no positive mutation calls were reported with FFPE wild type samples.

b) Limit of Detection (LoD)—RNA copies/µl

For the FGFR Kit, the limit of detecting FGFR alteration positive RNA in a background of wild-type RNA is defined as the lowest concentration for each FGFR alteration that is possible to detect with 95% probability. In this case, the LoD is reported as FGFR alteration positive RNA copies/µl. The LoD study was conducted in two steps due to limited clinical specimen availability. (1) IVTs to determine LoD; (2) use clinical specimens to verify the LoD.

LoD Determination:
Study was carried out to determine the LoD for each assay within the FGFR Kit using in vitro transcripts (IVT-RNA) samples. Individual mutant IVTs, one for each target, were spiked into a pool of normalized WT clinical RNA, serially diluted at levels above, at and below the LoD and tested. Approximately 60 technical replicates of each dilution point in the series were tested using three FGFR Kit lots (20 replicates per kit lot per dilution point). The LoD values for each assay was determined as the highest value (FGFR alteration positive RNA copies/µl) across all kit lots and confirmed using clinical samples for the following alterations: p.R248C (c.742C>T),

p.S249C (c.746C>G), p.G370C (c.1108G>T), p.Y373C (c.1118A>G), FGFR3-TACC3v3 and FGFR3-TACC3v1.

LoD Verification:
The LoD of the p.R248C (c.742C>T), p.Y373C (c.1118A>G), FGFR3-TACC3v3 and FGFR3-TACC3v1 assays was confirmed as part of the Contrived Sample Functional Characterization Study.

To confirm the LoD for the p.S249C (c.746C>G) and p.G370C (c.1108G>T) assays, RNA samples were extracted from FFPE UC tissues containing the p.S249C (c.746C>G) or p.G370C (c.1108G>T) alteration. The extracted RNA samples were normalized to an RNA input concentration of 16.67 ng/µl and screened with the FGFR Kit to determine the starting mutant CT value. These normalized mutant RNA samples were then spiked into normalized (16.67 ng/µl) WT FFPE UC samples to achieve the previously estimated LoD concentration. For the p.S249C (c.746C>G) and p.G370C (c.1108G>T) assays the estimated LoD was 289.82 and 141.57 copies/µl respectively. Twenty-four (24) replicates of both p.S249C (c.746C>G) and p.G370C (c.1108G>T) targets were tested at LoD using one lot of the FGFR Kit.

The LoD for the FGR2:BICC1, FGFR2-CASP7, or FGFR3:BAIAP2L21 targets could not be verified as Clinical samples could not be sourced due to their low prevalence (<1%).

The LoD values of the FGFR Kit assays are reported in Table 8.

Table 8: The LoD of each FGFR assay reported in terms of FGFR alteration positive RNA copies/µl

Reaction Mix	Alteration	Probability	LoD (RNA copies/µl)
Mut-1	p.R248C (c.742C>T)	0.95	75.80
	p.G370C (c.1108G>T)	0.95	141.57
Mut-2	p.S249C (c.746C>G)	0.95	289.82
	p.Y373C (c.1118A>G)	0.95	274.71
Fus-1	FGFR3-TACC3v3	0.95	45.75
	FGFR3-BAIAP2L1	0.95	9.07*
	FGFR2-CASP7	0.95	27.18*
Fus-2	FGFR3-TACC3v1	0.95	25.26
	FGFR2-BICC1	0.95	14.34*

* Due to the limited incidence of clinical UC cases harboring these alterations, the LoD of the FGFR3-BAIAP2L1, FGFR2-BICC1 and FGFR2-CASP7 assays was not verified with clinical specimens.

c) Limit of Detection (LoD)—Ct cut-offs

The C_T cut-off value for each assay in the FGFR Kit was determined using contrived sample dilution series. A total of eight dilutions per FGFR alteration were tested using three FGFR Kit lots (to generate 24 replicates per dilution level) over three days. These data were used to calculate false negative and false positive rates for proposed cut-off C_T values ranging from 20 to 45. The false negative and false positive rates were calculated with the contrived sample dilution which was estimated to be at or above the LoD for the respective assay. The final cut-off value for each assay was determined by assessing the expected false negative fractions and false positive fractions for each possible cut-off, for each alteration, in order to identify a value at which false positives and false negatives are unlikely to occur. A summary of these data is provided in Table 9. The determined LoD values are the sample cut-offs for target positivity. If a sample has a C_T lower or equal to this value it was determined as "FGFR Alteration Detected" for the corresponding alteration.

Table 9: Cut-off Values for each FGFR Alteration Assay

FGFR Assay	C_T Cut-off value
p.R248C (c.742C>T)	36.00
p.S249C (c.746C>G)	39.09
p.G370C (c.1108G>T)	41.00
p.Y373C (c.1118A>G)	43.00
FGFR3-TACC3v1	43.00
FGFR3-TACC3v3	43.00
FGFR3-BAIAP2L1	43.00
FGFR2-BICC1	43.00
FGFR2-CASP7	42.00

3. Linearity

The linearity of the positive controls (PC) was evaluated.

Three different PC lots of GMP-grade manufactured material were used. A ten-fold dilution series with six dilution levels, from 1.00E-03 ng/µL (Dilution 2) to 1.00E-09 ng/µL (Dilution 8), was created for each PC. Each PC sample was tested with three FGFR Kit lots. Three technical replicates of each dilution level and PC lot were assessed across each FGFR Kit lot, generating a total of 27 replicates per dilution level. A total of 18 RGQ runs were performed and submitted for statistical analysis.

The results showed that the amplification efficiency of all assays was similar, with a range between 93% and 107%, except for the Y373C assay, which had a lower efficiency (84.23%). This was also reflected in the regression plot with the Y373C

assay generating later Ct values across the dilution series (approximately a 2Ct shift compared to the results obtained with the remaining assays).

All PC lots tested generated similar results in the linear range of each of the assays. The linearity of the assay decreased with the increase of the dilution factor.

4. Analytical Specificity and Cross-reactivity

a) Analytical Specificity—*in silico* cross-reactivity

The level of potential cross-reactivity between primers, probes and blockers used within the FGFR Kit and non-FGFR targets within human and non-human genomes and the level of potential hetero-dimer was investigated. An in-silico analysis was performed to determine whether the primers, probes and blockers used with the FGFR Kit assays bind non-specifically within any genome, including the human genome. An additional in silico analysis was performed to determine whether the oligos used within each multiplex assay bind non-specifically to each other.

The in-silico analysis of oligo hetero-dimer indicated that there is a low probability of hetero-dimer formation. The primers and probes do not cross react with either WT alleles or any FGFR alterations not covered by the FGFR Kit and therefore will not cause a false signal.

In summary, in silico analysis of all primers and probes predicts specific amplification of the region around the target human FGFR2 and FGFR3 alteration sites with non-specific amplification from either human or tested microbial genomes unlikely.

b) Assay Cross-reactivity

Assay cross-reactivity is defined as non-specific amplification of an FGFR alteration utilized within the FGFR Kit other than the intended target for an assay, which gives a C_T value below the selected cut-off for that assay. Cross-reactivity was assessed using contrived specimens which had been produced to represent a high level for each alteration detected by the FGFR Kit.

A single contrived sample was created for each of the nine mutations. This represents high mutation level for each FGFR target. IVT's were spiked into normalized WT RNA to generate a stock of mutant positive contrived sample.

Samples were tested in triplicate with each of three FGFR Kit lots (nine replicates per sample). The proportion of false positive calls was assessed for each specimen tested with each assay. No amplification below the cut-

off was observed in any of the assays. It was concluded therefore that there was no cross-reactivity between the assays within the FGFR Kit.

5. Potential Interfering Substances

This study aimed to demonstrate the impact of potential endogenous and exogenous interferents on the FGFR alteration status detection across the RNeasy® DSP FFPE kit and the FGFR Kit.

Clinical WT UC and contrived samples representing the nine targets: p.R248C (c.742C>T), p.G370C (c.1108G>T), p.S249C (c.746C>G), p.Y373C (c.1118A>G), FGFR3-TACC3v3, FGFR3-BAIAP2L1, FGFR2-CASP7, FGFR3-TACC3v1 and FGFR2-BICC1 were used to test the effect of four interfering substances: Hemoglobin, Buffer RPE, Deparaffinization Solution and Paraffin Wax when tested with the FGFR System. Contrived samples were generated by using IVT-RNA spiked into normalized clinical WT at 3x LoD. These contrived samples were spiked with an interferent or interferent control either during the extraction phase or during sample normalization. The samples were extracted using the RNeasy DSP FFPE Kit and then were tested using the FGFR kit. Six replicates per samples were tested using the FGFR kit meaning a total of 60 replicates were tested per interferent and control (interferent-free sample) across the nine targets and WT.

To verify that an interferent has had no impact on assay performance for any sample, the following criteria must hold:

- The estimated difference between the test sample and the interferent-free control is not statistically significant (p-value >0.05) or, if the p-value is statistically significant (p-value ≤0.05), then;
- The absolute value of the estimated difference in Ct between the spiked interferent and the interferent-free control sample is less than 3x the intermediate precision of the assay as reported in the Repeatability and Reproducibility study.

Exogenous substances:
The testing concentrations of each exogenous interferent were estimated based on the highest (worst case) volume that could be carried over to the final sample eluate during the washing and purification steps in the RNeasy DSP FFPE Kit extraction process. These estimates take into consideration the maximum initial volume of each potential interferent, the step at which each interferent is introduced into the extraction process and the subsequent final volumes of each step where multiple reagents have been combined. The estimated maximum level of each interfering substance in the final eluate is summarized in Table 10 below.

Table 10: Estimated Maximum Level of Each Interfering Substance

Substance	Maximum Amount in the procedure (µl)	Estimated max volume in 30 (µl) eluate (µl)
Paraffin Wax	N/A	N/A
Deparaffinization Solution	480	0.019
Buffer RPE	1000	3

Endogenous substances (Hemoglobin):
A significant difference in mean Ct values was seen for alteration positive targets BAIAP2L1, CASP7, Y373C and R248C when tested with hemoglobin (BAIAP2L1) DPS (CASP7) and paraffin (Y373C and R248C). These samples showed a P-value of <0.05, however, the difference within the Ct values was still within the 3x intermediate precision for the assays. Therefore, all alteration positive samples passed the study acceptance criteria. For WT samples the IC passed the study acceptance criteria with all Ct values providing a p-value of >0.05 or the difference in means was within 3 x the intermediate precision of each of the assays. Additionally, all alteration positive and WT samples had a 100% rate of correct target calls for each PCR mix.

Six replicates per sample were tested using the FGFR Kit, making a total of 60 replicates per interferent and control. The four tested interferents showed no statistically significant difference between the control samples and test samples with the interferant added and did not cause an incorrect target calling of FGFR Alteration status.

6. Repeatability and Reproducibility

The within-site repeatability, intermediate precision and reproducibility of the FGFR Kit were investigated by testing contrived samples at 3x LoD, representing all nine alterations in the FGFR Kit, clinical specimens at 1xLoD representing common mutations, and a wild-type sample. These RNA samples were divided into single use aliquots for testing with the FGFR Kit at three independent laboratories (one site in the UK and two sites in the USA).

To demonstrate repeatability, contrived specimens (IVTs) at 3x LoD and wild-type samples were tested at QIAGEN Manchester by three operators over ten days using three RGQ instruments to generate a total of 60 replicates per FGFR alteration. The acceptance criteria stated that for the wild-type and 3x LoD FGFR alteration contrived samples (separately), the lower two-sided exact 95% confidence limit (CL) for the overall percentage of correct calls must be ≥90%. All samples met these acceptance criteria. The proportion of correct calls for each sample tested at Site 1 (repeatability) along with the corresponding two-sided exact 95% CLs are reported in Table 11.

Table 11: Repeatability - Proportion of Correct Results based on Overall FGFR Alteration Status, (i.e. "FGFR Alteration Detected" or "No Alteration Detected" for each Assay

Reaction Mix	Template	Fraction	Percentage	Lower Two-Sided 95% CL	Upper Two-Sided 95% CL
Mut-1	p.R248C (c.742C>T)	60 / 60	100.00%	94.04%	100.00%
	p.G370C (c.1108G>T)	60 / 60	100.00%	94.04%	100.00%
Mut-2	p.S249C (c.746C>G)	60 / 60	100.00%	94.04%	100.00%
	p.Y373C (c.1118A>G)	60 / 60	100.00%	94.04%	100.00%
Fus-1	FGFR3-TACC3v3	59 / 60	98.33%	91.06%	99.96%
	FGFR3-BAIAP2L1	60 / 60	100.00%	94.04%	100.00%
	FGFR2-CASP7	60 / 60	100.00%	94.04%	100.00%
Fus-2	FGFR3-TACC3v1	59 / 60	98.33%	91.06%	99.96%
	FGFR2-BICC1	60 / 60	100.00%	94.04%	100.00%
Mut-1	wild-type	60 / 60	100.00%	94.04%	100.00%
Mut-2		60 / 60	100.00%	94.04%	100.00%
Fus-1		60 / 60	100.00%	94.04%	100.00%
Fus-2		59 / 60	98.33%	91.06%	99.96%

Reproducibility was measured by testing contrived samples at 3× LoD level, clinical samples close to LoD and wild-type samples across the three different sites (1 internal Qiagen site in the United Kingdom and two additional external sites in the USA). The contrived samples for all FGFR alterations at 3×LoD and wild-type samples were tested by three operators (per site) over five days using three RGQs at each external site. In addition, RNA extracted from formalin-fixed paraffin embedded (FFPE) Urothelial Cancer (UC) clinical samples was used to test the device reproducibility. Testing was performed at the LoD level for each target (p.R248C (c.742C>T), p.S249C (c.746C>G), p.G370C (c.1108G>T), p.Y373C (c.1118A>G), FGFR3-TACC3v1 and FGFR3-TACC3v3) utilizing clinical samples. Clinical samples could not be sourced for BAIAP2L1, BICC1 or CASP7. The analysis also combined lot to lot variability within the study design.

Samples were extracted, pooled and normalized to 16.67 ng/µl. The normalized FGFR alteration positive samples were then diluted to the LoD of the respective assay in a background of WT RNA. All 1x LoD clinical samples were tested daily at each of the three sites using the FGFR Kit (2 biological replicates x 2 FGFR Kit lots x 2 operator x 3 days= 24 replicates at each site). These total replicates were tested across three RGQ instruments at each site and 2 of the 3 kit lots were used alternatively in each site.

The proportion of correct calls for each sample, along with the corresponding two-sided exact 95% CLs, are reported for all sites (reproducibility) is reported in Table 12. To determine the intermediate precision of each FGFR Kit alteration assay, a random effects model was fitted to the data with C_T value as the response variable and day, operator, RGQ and run as categorical random effects. The results met the acceptance criteria.

Table 127: Assay reproducibility – Number of correct calls and two-sided 95% confidence limits for each FGFR alteration and wild-type samples tested across all sites

Mutation	Target level	Specimen type	% Agreement (No. positive results/total)	two-sided 95% confidence limit
p.R248C (c.742C>T)	3xLoD	Contrived	100.00% (120/120)	96.97%, 100.00%
	1xLoD	Clinical	100.00% (72/72)	95.01%, 100.00%
p.G370C (c.1108G>T)	3xLoD	Contrived	100.00% (120/120)	96.97%, 100.00%
	1xLoD	Clinical	98.61% (71/72)	92.50%, 99.96%
p.S249C (c.746C>G)	3xLoD	Contrived	100.00% (120/120)	96.97%, 100.00%
	1xLoD	Clinical	100.00% (72/72)	95.01%, 100.00%
p.Y373C (c.1118A>G)	3xLoD	Contrived	100.00% (120/120)	96.97%, 100.00%
	1xLoD	Clinical	98.61% (71/72)	92.50%, 99.96%
FGFR3-TACC3v3	3xLoD	Contrived	99.17% (119/120)	95.44%, 99.98%
	1xLoD	Clinical	98.61% (71/72)	92.50%, 99.96%
FGFR3-TACC3v1	1xLoD	Contrived	99.17% (119/120)	95.44%, 99.98%
	⅓xLoD	Clinical	87.50% (63/72)	77.59%, 94.12%
FGFR3-BAIAP2L1	3xLoD	Contrived	100.00% (120/120)	96.97%, 100.00%
	1xLoD	Clinical	N/A*	
FGFR2-BICC1	3xLoD	Contrived	100.00% (120/120)	96.97%, 100.00%
	1xLoD	Clinical	N/A*	
FGFR2-CASP7	3xLoD	Contrived	100.00% (120/120)	96.97%, 100.00%
	1xLoD	Clinical	N/A*	
WT (Mut-1 Reaction Mix)	N/A	Clinical	100.00% (120/120)	96.97%, 100.00%
WT (Mut-2 Reaction Mix)	N/A	Clinical	100.00% (120/120)	96.97%, 100.00%
WT (Fus-1 Reaction Mix)	N/A	Clinical	100.00% (120/120)	96.97%, 100.00%
WT (Fus-2 Reaction Mix)	N/A	Clinical	96.67% (116/120)	91.69%, 99.08%

*Clinical samples could not be sourced for BAIAP2L1, BICC1 or CASP7.

Refer to Section IX.C. Additional Studies for a description of additional reproducibility data that is provided postmarket.

7. **Specimen Handling—reproducibility**

To demonstrate that different laboratories will produce acceptable results when starting from the same FFPE UC specimen, extractions were performed across three different sites. Clinical wild-type FFPE specimens, and four FGFR alteration detected clinical FFPE specimens, each containing one of the p.G370C (c.1108G>T), p.S249C (c.746C>G), FGFR3-TACC3v1 or FGFR3-TACC3v3 alterations. These alterations were selected as representatives for each of the four reaction mixes in the FGFR Kit. Specimens were divided into three independent sets and extracted at the three sites. Each set was extracted three times by two at each site. In total, 185 replicates were tested across all samples. When comparing the results of each sample across all three sites, the percentage of correct alteration calls for alteration detected, and alteration not detected samples was 96.22% (178

of 185 samples tested). Therefore, when the same sample is extracted at different laboratories the proportion of concordant calls was greater than 95% demonstrating an acceptable comparable performance between operators and different laboratories.

8. Lot-to-Lot Interchangeability

This study was to demonstrate lot interchangeability and consistency of FGFR alteration detection by the FGFR System. The FGFR System utilizes two separate kits: (1) the RNeasy® DSP FFPE Kit for isolation of RNA from FFPE UC specimens, and (2) the FGFR Kit for the amplification and detection of the isolated RNA for its FGFR alteration status. A combination of clinical and contrived samples was used for this study. Four clinical FFPE specimens harboring the p.R248C (c.742C>T), p.S249C (c.746C>G), FGFR3-TACC3v1 alterations and a clinical wild-type FFPE specimen were extracted in duplicate using three RNeasy® DSP FFPE Kit lots and tested with three different FGFR Kit lots to generate a total of 36 replicates for each sample. All RNA samples were normalized and diluted to the 3x LoD level for each alteration. In addition, contrived samples for the p.G370C (c.1108G>T), p.Y373C (c.1118A>G), FGFR3-TACC3v3, FGFR3-BAIA2PL1, FGFR2-CASP7 and FGFR2-BICC1 alteration were produced at the 3x LoD level and tested using the three FGFR Kit lots. Contrived samples were generated to the 3x LoD level with normalized RNA extracted from clinical FFPE wild-type specimens using one lot of RNeasy® DSP FFPE Kit. The test results were analyzed for all kit lots and for each FGFR alteration assay. The overall percentage of correct calls for all samples across all FGFR Kit and RNeasy® DSP FFPE Kit lots was 99.65% (286 of 287 samples tested), demonstrating that combining different lots of the RNeasy® DSP FFPE Kit and FGFR Kit does not impact the ability of the FGFR System to determine a correct FGFR alteration call.

9. Contrived Sample Functional Characterization Study

To support the use of in vitro transcript (IVT) as an alternative for clinical samples for certain variants, which are difficult to acquire, a study was conducted to evaluate the functional behavior of the contrived samples to that of clinical specimens using the FGFR Kit by comparing the performance of the Contrived model using individual FGFR alteration positive IVTs spiked in a background of WT clinical RNA derived from FFPE tissue from patients with advanced UC, and FGFR alteration positive Clinical RNA derived from FFPE tissue from patients with advanced UC, for representative targets within each of the FGFR multiplex reaction mixes.

Contrived samples harboring the *p.R248C (c.742C>T), p.Y373C (c.1118A>G), FGFR3:TACCv3 and FGFR3:TACCv1* alterations were prepared. Samples were produced to levels around, encompassing and below the estimated LoD, based on IVT copy number. RNA extracted from procured clinical samples harboring the selected mutations above were tested with the FGFR Kit in order to generate a CT value for each sample. Extracted samples were pooled and normalized to attain 250 ng of RNA input into the RT reaction. Samples were diluted (based on CRTR value)

using RNA extracted from clinical samples and WT for all the target genotypes, to levels around, encompassing and below the estimated LoD.

The performance of Contrived and Clinical samples was determined at levels around the estimated LoD by comparing positivity rates for each of the sample sets. In order to demonstrate equivalence, following parameters were considered:

- Probit Model Selection: positivity rates were used to evaluate three Probit models for each sample set across both Clinical and Contrived sample types
- Significance (at 5% level) of interaction term and sample type term
- Significance (at 5% level) of differences in positivity rates
- Difference (absolute and fold difference) in estimates of C95, C75 and C50 (if applicable)

Three Probit models were evaluated for each sample set across both sample types (contrived, clinical) to assess the relationship between positivity rate [number of amplifications / (number of amplifications + number of non-amplifications)] and sample type. The likelihood ratio test was used to compare the models.

Based on the positivity rates and the Probit models of the four representative sample sets tested (p.R248C (c.742C>T), p.Y373C (c.1118A>G), FGFR3-TACC3v3 and FGFR3-TACC3v1), the proposed contrived specimen (IVT) has a similar performance when compared to the clinical samples. Although the values at C95 for two fusions seems to have large difference between IVT and clinical specimen, overall, the difference between these two specimen types may not make impact on clinical significance.

10. Guard Band Studies

Guard band studies were designed to demonstrate that the FGFR Kit is robust in performance to changes that may be introduced into the FGFR Kit workflow by the end user. For all guard band studies, in order for a test condition to be deemed to have no impact on assay performance the difference in mean C_T between the specified test condition and the nominal condition for each sample type must be within +/-2x SD of the assay.

a) Volumetric Guard Band

This study determined the effect of varying reagent volume on the alteration status of samples called by the FGFR Kit. The volumetric tolerance was tested by varying the volume of each individual component in the reverse transcription and PCR reactions while keeping the volume of the other components constant. Each component volume was varied by ±3% and ±6%. This represents the total error that can be introduced by pipetting calculated by relative accuracy and precision stated in the pipette specifications.

A total of 33 conditions were tested for the reverse transcription step and 16 different conditions were tested for the PCR step, using contrived samples for the p.R248C (c.742C>T), p.Y373C (c.1118A>G), FGFR3-TACC3v1 and FGFR3-TACC3v3 alterations at 3x LoD of the assays. These samples were considered to be representative of each of the four multiplex reaction mixes and the green and orange RGQ channels. For each condition, tube, alteration and sample type the proportion of correct calls was reported along with the corresponding two-sided 95% CI.

The impact of differences in pipetting volumes was determined from the mean C_T values generated for all assays when tested with the variable conditions compared to the nominal conditions. Three reaction mixes (Mut-1, Fus-1 and Fus-2) met the study acceptance criteria. While the pre-specified criteria were not met for the Mut-2 reaction mix where reverse transcription components were varied by ±6%, the predicted C_T value was within the sample acceptance criteria for the p.Y373C (c.1118A>G) assay and therefore did not impact the sample alteration call at the 3x LoD level. This range is outside of the expected variation (±2%) that would be introduced when using calibrated pipettes according to manufacturer's recommendations.

b) Mixing Guard Band

This study aimed to determine the impact of varying the mixing method for reagents and sample in the reverse transcription and PCR steps of the FGFR workflow on the performance of the FGFR Kit, and to define the recommended mixing method for the reverse transcription and PCR steps.

A combination of inversion and vortex mixing conditions were tested at the reverse transcription and PCR steps using four contrived samples (p.R248C (c.742C>T), p.Y373C (c.1118A>G), FGFR3-TACC3v1, and FGFR3-TACC3v3) representative of each of the FGFR reaction mixes. Twelve (12) replicates per sample were generated for each mixing condition tested using the respective PCR reaction mix. All samples and test conditions met the study acceptance criteria and variations in mixing procedure between inversion and vortex mixing had no significant effect on the representative alterations tested. The final recommended mixing method for the FGFR Kit is vortexing.

c) Reverse Transcription Incubation Guard Band

The FGFR Kit requires a suitably calibrated incubation system capable of incubating at 42ºC and 95ºC. The objective of this study was to determine the tolerance of the FGFR Kit to possible variations of ±1.5ºC (i.e. 40.5ºC, 42ºC, 43.5ºC) during the incubation step of the reverse transcription reaction. Contrived samples for each of the four representative FGFR Kit

alterations (p.R248C (c.742C>T), p.Y373C (c.1118A>G), FGFR3-TACC3v1 and FGFR3-TACC3v3) were produced by spiking normalized wild-type RNA with single FGFR alteration detected IVTs. Wild-type samples were assessed with all four reaction mixes of the FGFR Kit and the four representative contrived samples were tested with the corresponding reaction mix for that alteration. Ten (10) replicates per contrived sample were generated for each reverse transcription incubation condition, seven replicates of wild-type sample were tested. For all samples tested, the difference in mean C_T between the test and nominal conditions were within ±2x SD of the assay. All assays reported 100% correct FGFR alteration calls across all conditions and samples with two exceptions:

1. The percentage of correct calls for the FGFR3-TACC3v1 sample tested with the Fus-2 reaction mix tested at 43.5˚C was 90% (9/10).
2. The percentage of correct calls for the wild-type sample tested with the Mut-2 reaction mix tested at 43.5˚C was 90% (9/10).

The difference in mean C_T between the 43.5˚C and 42˚C (the nominal condition) for the FGFR3-TACC3v1 sample (tested with the Fus-2 reaction mix) and the wild-type sample (tested with the Mut-2 reaction mix) was 0.0208 and 0.050, respectively. These differences and the corresponding two-sided 95% CIs were within twice the standard deviation of the assay, demonstrating that variations in the temperature of the reverse transcription step by ±1.5ºC do not impact the C_T values generated by the FGFR assay.

d) <u>PCR cycling Guard Band</u>

This study was performed to characterize the robustness of the FGFR Kit to variations in RGQ temperature during the annealing step of the PCR reaction. The RGQ instrument has set specifications for PCR cycle temperature. However, alterations in RGQ instrument performance may result in different annealing temperatures which may impact assay performance. The calibration range of the RGQ instrument is ±0.5°C of the target temperature for the rotor and reaction volume used for the FGFR PCR reactions.

The thermal cycling profile was guard banded by varying the annealing temperature by ±0.5°C, ±1°C and ±2°C from the nominal annealing temperature (60°C). Wild-type RNA and contrived samples at 3x LoD level were tested for all nine alterations detected by the FGFR Kit. Ten replicates were tested per sample and test condition.

There was no effect of variations in annealing temperatures within the calibration range of the RGQ instrument for all assays except p.Y373C (c.1118A>G). Additional testing showed that p.Y373C (c.1118A>G) samples diluted just under the LoD for the assay had a hit rate (FGFR

alteration positive rate) of 96.15% when tested on the lower edge of the calibration range, demonstrating that the shift in CT values did not impact the FGFR alteration call at or above the LoD for that assay. These results showed that the FGFR Kit is able to tolerate variations in the PCR annealing temperature which are within the calibration range for the RGQ instrument.

e) Freeze-thaw Cycles Guard Band

The objective of this study was to determine the tolerance of the FGFR Kit to up to six freeze-thaw cycles and to determine the number of freeze-thaw cycles (in-use stability) that can be introduced by the end user. Both the reverse transcription and PCR steps of the workflow were tested in this study. Contrived samples for each of the four representative FGFR Kit alterations (p.R248C (c.742C>T), p.Y373C (c.1118A>G), FGFR3-TACC3v1 and FGFR3-TACC3v3) were produced at 3x LoD for each assay and tested to generate nine to ten replicates per sample per test condition over ten PCR runs in total. All differences in mean C_T values between the nominal and test conditions met the study acceptance criteria. Based on these data, a maximum of five freeze-thaw cycles are recommended for the FGFR Kit.

f) Thawing and Set-up Time Guard Band

The effect of varying thawing and set-up time of the FGFR Kit reagents and samples on FGFR alteration status was evaluated. The robustness of both the reverse transcription and PCR steps of the workflow was investigated. A matrix experiment was designed with a total of three conditions for thawing time and two conditions for set-up time to be investigated. A maximum set-up time of four hours and a thawing time of three hours were tested. The nominal condition in the study was defined as a thawing time of 1 hour and a run set-up time of 1 hour. Contrived samples for each of the four representative FGFR Kit alterations (p.R248C (c.742C>T), p.Y373C (c.1118A>G), FGFR3-TACC3v1 and FGFR3-TACC3v3) were produced at 3x LoD for each assay and tested to generate ten replicates per sample per test condition. All results obtained were within 2x SD of each assay, indicating that the FGFR Kit is robust to thawing times up to three hours at room temperature and set-up times of up to four hours.

g) RNA Input Guard Band

The objective of this study was to demonstrate the robustness of the kit to errors in measurement of RNA concentration and/or variation in the amount of degraded RNA was tested. A range of RNA inputs from 10% (25ng) and 200% (500ng) of the recommended 250 ng RNA input was tested. This range is considered to be outside the expected variance which may be introduced by RNA quantification errors and/or variation in the amount of

degraded RNA during routine use of the *therascreen* FGFR RGQ RT-PCR Kit.

RNA was extracted from FGFR alteration positive Clinical specimens (p.S249C (c.746C>G) and FGFR3-TACC3v1) and targeted to the 3xLoD level in a background of pooled RNA extracted from wild type Clinical specimens at four different RNA inputs: 500ng, 300ng, 250ng (standard input) and 25ng. Twenty replicates of each sample were tested using one *therascreen* FGFR RGQ RT-PCR Kit lot.

All 80 replicates of S249C samples at the four RNA input levels tested gave the correct alteration call. For FGFR3-TACC3v1, the correct alteration call was obtained for RNA input levels of between 25 and 300ng (59/60 replicates), demonstrating that RNA input levels within this concentration range doesn't impact the performance of the Kit. Two of 20 replicates tested at the 500ng RNA input level gave an incorrect call. It can therefore be concluded that the *therascreen* FGFR RGQ RT-PCR Kit performs as expected at RNA input levels ranging from 25 – 300 ng.

h) Proteinase K Guard Band

This study assessed the effect of variation in Proteinase K incubation temperatures during the RNA extraction procedure on the ability of the *therascreen* FGFR RGQ RT-PCR Kit to correctly call the FGFR alteration status of clinical samples.

Three clinical samples, two wild-type and one Y373C alteration positive sample, were extracted using one RNeasy FFPE Kit lot across nine test conditions, where the incubation temperatures for the Proteinase K activation and deactivation were each altered across a ± 5°C range. Extracted samples were then tested with one *therascreen* FGFR RGQ RT-PCR Kit lot. The *therascreen* FGFR RGQ RT-PCR Kit was able to consistently and accurately call the Y373C positive samples in 100% (108/108) of cases, and wild-type samples in 97.5% (117/120) of cases, across the different extraction conditions tested.

These data show that variations of up to ±5°C in the incubation temperature of Proteinase K activation and deactivation during the RNA extraction procedure has no adverse effect on the ability of the *therascreen* FGFR RGQ RT-PCR Kit to correctly call the FGFR alteration status of a sample.

11. Cross-Contamination/Analytical Carryover

The study examined the entire therascreen® FGFR RGQ RT-PCR System from extraction to PCR runs and investigated if carryover occurred between samples, extractions, and within or between runs when high FGFR alteration detected samples are tested adjacent to FGFR alteration not detected (wild-type) samples.

Clinical FFPE UC specimens were divided into two independent sets. Both sets comprised eighteen wild-type FFPE specimens, with an additional six p.S249C (c.746C>G) detected FFPE specimens (Set A) or six FGFR3-TACC3v1 detected FFPE specimens (Set B). The RNA extraction, reverse transcription and PCR reaction setup followed a matrix designed to introduce the risk of cross-sample contamination by testing high levels of FGFR alteration samples adjacent to wild-type samples at each step in the workflow. Each set was tested by a different operator using the same FGFR Kit lot. Four runs (four reverse transcription runs plus four real-time PCR runs) were performed by each operator, making eight runs in total. Only Mut-2 (Set A) and Fus-2 (Set B) reaction mixes were tested. A total of 128 wild-type replicates were tested. In order to meet the study acceptance criteria, the percentage of false positive calls for wild-type samples should be less than 5%. The percentage of false positive calls for wild-type samples was 3.13% (4 of 128 samples). These results confirmed that when following the *therascreen*® FGFR RGQ RT-PCR System workflow, the cross-contamination rate of wild-type samples by FGFR alteration detected samples that share the RNA extraction and run set-up procedure is less than 5%.

12. <u>Specimen Stability</u>

The purpose of this study was to establish a storage duration for extracted RNA, to determine a number of freeze-thaw cycles allowed for extracted RNA and to provide a storage duration at specified temperatures for synthesized cDNA. For this study, Clinical FFPE UC specimens harboring the p.R248C (c.742C>T), p.Y373C (c.1118A>G), FGFR3-TACC3v1 and FGFR:TACC3v3 alterations, and a clinical FFPE UC specimen which was wild-type for the FGFR alterations of interest were extracted using the RNeasy® DSP FFPE Kit.

The stability of the normalized RNA was demonstrated by storing extracted RNA (which had been dispensed in to single use aliquots) at -100°C to -65°C for up to 34 days and tested at five different time points. The maximum number of freeze - thaw cycles the normalized RNA samples can withstand was established by storing the extracted RNA at -100°C to -65°C for up to 34 days. The same aliquot of RNA (five freeze-thaw cycles) was tested at five different time points. The stability of the synthesized cDNA samples was demonstrated by storing closed single use aliquots of a pool of cDNA under different conditions (-30°C to -15°C and 2°C to 8°C) and testing them at different time points. For each sample, reaction mix and alteration tested the allowable drift limit was defined as the baseline mean (calculated at Testing Time Point Zero) ±2x the corresponding standard deviation of the assay. Based on the results of this study, recommended storage times for extracted RNA and cDNA samples, and the maximum number of freeze-thaw cycles recommended for extracted RNA for use with the FGFR Kit were determined.

RNA was extracted from FFPE material and testing at five different time points. The samples were extracted using the RNeasy® DSP FFPE kit at each time point. For each sample, reaction mix and alteration tested the allowable drift limit was defined as the baseline mean (calculated at Testing Time Point Zero) ±2x the corresponding standard deviation of the assay. The acceptance criteria for this study was that the stability limit of the samples would be determined as the point for which the regression line of the corresponding target C_T value against time intersects with the pre-defined allowable drift limit. This study demonstrated that the RNA in clinical FFPE sections mounted on un-stained glass slides for representative samples of wild-type, p.R248C (c.742C>T), p.Y373C (c.1118A>G), FGFR3-TACC3v1 and FGFR3-TACC3v3 are stable for at least four weeks from the first tested time point.

13. Reagent Stability

Testing of the stability of the FGFR Kit under possible usage conditions is on-going. The stability study consists of the following elements:

- Real-time stability (closed bottle, post-transport simulation): demonstrates stability of the kit under intended storage conditions.
- In-use stability (including freeze-thaw and open vial, post-transport simulation): demonstrates stability of kit reagents under simulated usage conditions including stability of reagents that have been opened and returned to storage conditions.
- Transport simulation study (integrated in real-time and in-use stability study): demonstrates stability of reagents that have been handled under simulated shipping conditions.

For each part, reagents are stored under prescribed conditions and then used to perform testing using the FGFR Kit. To demonstrate stability, the study acceptance criteria must be met.

B. Animal Studies

None.

C. Additional Studies

As pat of the approval with conditions, the following data will be obtained:

1. Obtain clinical validation data for FGFR2 fusions (FGFR2:BICC1 and FGFR2:CASP7) and FGFR3 fusion (FGFR3:BAIAP2L1) from the ongoing phase 3 clinical trial and update the labeling;
2. Obtain additional analytical validation data for FGFR2 fusions (FGFR2:BICC1 and FGFR2:CASP7) and FGFR3 fusion (FGFR3:BAIAP2L1) using clinical specimens. These studies include analytical accuracy study, reproducibility study, and Limit of Detection study. The result from these studies should be included in the labeling.

3. Conduct a reproducibility study with specimens at LoD level for each of the 9 alterations according to protocol and Instructions For Use with the resolved NTC issue.
4. Provide results and documentation from regression testing on the commercial release configuration (*therascreen*® FGFR FFPE MDx Assay Profile v1.0.0) to confirm the correction to the fusion nomenclature.

X. SUMMARY OF PRIMARY CLINICAL STUDIES

The safety and effectiveness of the *therascreen* FGFR RGQ RT-PCR kit was evaluated in a retrospective study designed to demonstrate that the *therascreen* FGFR RGQ RT-PCR kit correctly detects the presence of genetic alterations in FGFR genes in urothelial cancer patients for the purpose of clinically validating the use of the test companion diagnostic test for BALVERSA™ (erdafitinib). Banked FFPE UC tissue samples collected during the Janssen Research & Development study BLC2001 (NCT02365597) were tested with the *therascreen* FGFR RGQ RT-PCR kit. The BLC2001 study assessed the safety and effectiveness of erdafitinib in a multicenter, open-label, single-arm study in patients with locally advanced or metastatic urothelial carcinoma who are eligible for treatment with erdafitinib on the basis of positive results for four point mutations (p.R248C (c.742C>T), p.S249C (c.746C>G), p.G370C (c.1108G>T) and p.Y373C (c.1118A>G)) and two fusions (FGFR3-TACC3v1 and FGFR3-TACC3v3) in the FGFR3 gene. The test was also designed to identify FGFR2 fusions (FGFR2-BICC1 and FGFR2-CASP7) and FGFR3 fusion FGFR3-BAIAP2L1 because these FGFR fusion patients were eligible for the trial, however, the Qiagen test is not clinically validated to detect these 3 fusions due to the lack of required clinical sample materials.

A. Study Design

Study 42756493BLC2001 (NCT02365597) was a multicenter, open-label, single-arm study to evaluate the efficacy and safety of BALVERSA (erdafitinib) in patients with locally advanced or metastatic urothelial carcinoma. A total of 210 patients were enrolled in three dose regimens. Ninety-nine patients were enrolled in Regimen 3, the selected regimen (8 mg daily with possible pharmacodynamically-guided uptitration to 9mg). Of these, eighty-seven patients had disease that had progressed on or after at least one prior chemotherapy (chemo-relapsed/refractory) and that had at least 1 of the following genetic alterations: FGFR3 gene mutations (R248C, S249C, G370C, Y373C) or FGFR gene fusions (FGFR3-TACC3, FGFR3-BAIAP2L1, FGFR2-BICC1, FGFR2-CASP7), as determined by a clinical trial assay performed at a central laboratory.

The primary objective of this study was to evaluate the objective response rate (complete response [CR]+ partial response [PR]) of the selected dose regimen per RECIST version 1.1 and blinded independent review committee (BIRC). The study was comprised of a Screening Phase, a Treatment Phase comprised of 28-day treatment cycles and a post-treatment Follow-up Phase that will extend from the End-of-Treatment Visit until the participant has died, withdraws consent, is lost to follow-up, or the end of the study.

The end of study is defined as approximately 12 months after last participant is enrolled. Safety was monitored throughout the study.

The 42756493BLC2001 study was sponsored by Janssen Research & Development and the data were submitted to the FDA in New Drug Application (NDA 212018) in support of the proposed indication.

Patients were screened for enrollment into this trial using an investigational clinical trial assay (CTA) designed to detect the specific FGFR alterations. Enrollment commenced in May 2015 and screening closure for Clinical Study Report was March 15, 2018. The Bridging study was conducted to establish the analytical and clinical concordance between the FGFR Kit and the CTA to clinically validate the test as safe and effective for selecting patients who may benefit with BALVERSA™ (erdafitinib).

The Bridging Study also included representativeness analysis to provide evidence that the study subjects and re-tested specimens represented the Device Intended Use population.

1. <u>Inclusion and Exclusion Criteria for Specimen CTA Testing</u>

Archival or freshly obtained FFPE tissue blocks were required for the assessment of FGFR alteration status to determine trial molecular eligibility. A minimum of 10% viable tumor cells, in relation to non-neoplastic cells within the tumor area was required for testing was mandatory. Tissue with a tumor area of $\geq$1 cm^2 (100 mm^2) per tissue section was recommended. If sites were unable to send a tissue block, a total of 15 tissue slides (4 microns) were submitted. If archived biopsy tissue was not available, an FFPE core biopsy from a metastatic site was required.

One section was stained with H&E stain to determine tumor presence and subject to pathology review to confirm the presence of 10% viable tumor cells, in relation to non-neoplastic cells within the tumor area. If the sample passed pathology review, the tissue samples were macrodissected and total RNA was extracted using the QIAGEN AllPrep DNA/RNA FFPE Kit.

2. <u>Clinical Inclusion and Exclusion Criteria for Patient Enrollment</u>

Enrollment in the 42756493BLC2001 study was limited to patients who met the following key inclusion and exclusion criteria:

Inclusion Criteria:

- Must have histologic demonstration of metastatic or surgically unresectable urothelial cancer.
- Must have measurable disease according to the Response Evaluation Criteria in Solid Tumors (RECIST, version 1.1) at baseline
- Must have an Eastern Cooperative Oncology Group (ECOG) performance status score 0, 1, or 2

- Must have adequate bone marrow, liver, and renal function as described in protocol
- Negative pregnancy test (urine or serum beta human chorionic gonadotropin [b-hCG]) at Screening for women of child bearing potential who are sexually active
- Must have shown disease progression according to RECIST, version 1.1, following prior chemotherapy for metastatic or surgically unresectable urothelial cancer

Exclusion Criteria:

- Received chemotherapy, targeted therapies, definitive radiotherapy, or treatment with an investigational anticancer agent within 2 weeks before the first administration of study drug.
- Has persistent phosphate level greater than upper limit of normal (ULN) during screening (within 14 days of treatment and prior to Cycle 1 Day 1) and despite medical management
- Has a history of or current uncontrolled cardiovascular disease
- Females who are pregnant, breast-feeding, or planning to become pregnant within 3 months after the last dose of study drug and males who plan to father a child while enrolled in this study or within 5 months after the last dose of study drug

Has not recovered from reversible toxicity of prior anticancer therapy

3. Follow-up Schedule

Subjects were to be treated until disease progression, unacceptable toxicity, or any other protocol-defined reason for treatment discontinuation. Safety evaluations included adverse event monitoring, physical examinations, electrocardiogram (ECG) monitoring, clinical laboratory parameters (hematology and chemistry), vital sign measurements, ECOG performance status, and ophthalmologic examinations. Blood samples were drawn for assessment of PK, clinical laboratory tests, biomarkers and pharmacodynamics parameters.

4. Clinical Endpoints

The primary objective was to evaluate the objective response rate (ORR; complete response [CR] + partial response [PR]) of the selected dose regimen of BALVERSA™ (erdafitinib) in subjects with metastatic or surgically unresectable UCs that have specific FGFR genomic alterations.

Response was assessed per RECIST version 1.1 and BIRC for all subjects.

B. Accountability of PMA Cohort

A total of 2214 patients were screened for 42756493BLC2001 trial using the CTA. Of those 1987 produced valid test results, of which 417 were FGFR alteration positive. Of

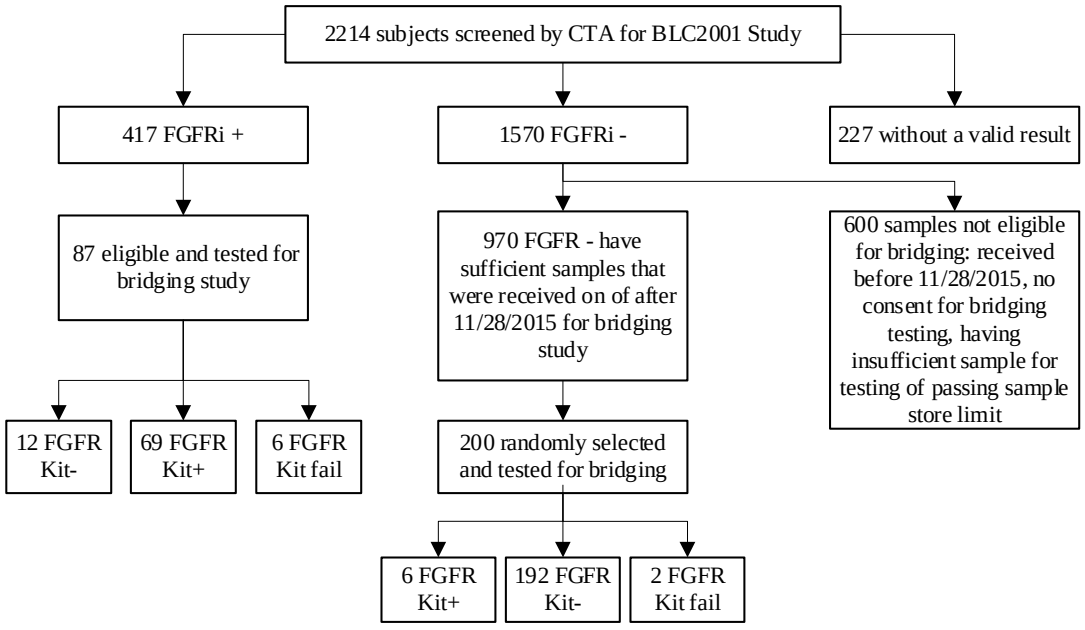
the 417 FGFR alteration positive patients, 87 chemo-relapsed/refractory patients were eligible and tested for the bridging study.

Additionally, 970 out of 1570 CTA negative patients were eligible for the bridging study and 200 of them were randomly selected for testing.

Eighteen specimens from patients with CTA alteration positive results did not have sufficient tissue remaining to perform RNA extraction for the Bridging Study. Therefore, residual CTA-derived RNA from the 42756493BLC2001 study was used for re-testing with the FGFR Kit.

Figure 1 summarizes the number and distribution of specimens from patients screened in the 42756493BLC2001 study and those re-tested by the FGFR Kit.

Figure 1: 42756493BLC2001 Study Screen and Bridging Study Sample Selection



C. Study Population Demographics and Baseline Parameters

1. Representativeness Analysis

To assess representativeness of the actual subjects in the Bridging Study versus the eligible subjects, it was initially intended to compare patient demographics and tumor characteristics of subjects

- with FGFR Kit evaluable (valid FGFR Kit result) with FGFR Kit not evaluable and
- FGFR Kit testable (sufficient material for testing in the Bridging Study and have consented for re-testing) and FGFR Kit not testable patients

However, due to the limitation of desired source data, and the difference between actual data collection and the definition of testing populations,

- the population subjected to molecular screening using the CTA (n=2214) were compared with those from the Bridging Study eligible subjects (n=1057)
- the 287 subjects tested in the Bridging Study were compared with the rest of subjects who are eligible for the Bridging Study but not selected for the testing (n=770)

P-values were calculated based on two-group t-test for continuous measures (Gender, Race, Ethnicity and Site of Tissue Collection) and Pearson chi-square test for category data (Age, % tumor area and % Viable cells). The results are reported in Table 13.

The demographics profile (age, gender, race and ethnicity) of subjects who were eligible for the Bridging Study (n=1057) is similar to the demographic profile of the CTA screened population (n=2214), indicating that the defined bridging eligible population represent the CTA screened population. The comparison of the demographic profiles between subjects tested and not tested in the Bridging Study further suggests that the selected subjects represent the FGFR Kit intended use population, recognizing that the p-values for ethnicity and percent viable cells are less than 0.05. Given the relatively large sample sizes for these analyses, it is not surprising to observe a statistical difference in some demographics. However, the differences with statistical significance are quite small and thus clinically insignificant.

Table 13: Demographic Data for Population Representativeness

	BLC2001 Screened n=2214	Bridging Eligible n=1057	Bridging Tested n=287	Not Tested n=770	p Value (Tested vs. Not Tested)
Age (years)					
Mean (SD)	66.6 (9.87)	66.6 (9.94)	66.7 (9.68)	66.6 (10.04)	0.87
Median	67.0	68.0	68.0	67.5	
Range	(19; 90)	(28; 90)	(31; 89)	(28; 90)	
Gender, n (%)					
Male	1687 (76.2%)	797 (75.4%)	217 (75.6%)	580 (75.3%)	0.92
Female	527 (23.8%)	260 (24.6%)	70 (24.4%)	190 (24.7%)	
Race, n (%)					
White	1454 (65.7%)	738 (69.8%)	193 (67.2%)	545 (70.8%)	0.08
Black	23 (1.0%)	10 (0.9%)	6 (2.1%)	4 (0.5%)	
Asian	312 (14.1%)	106 (10.0%)	27 (9.4%)	79 (10.3%)	
Other	425 (19.2%)	203 (19.2%)	61 (21.3%)	142 (18.4%)	
Ethnicity, n (%)					
Hispanic/Latino	35 (1.6%)	17 (1.6%)	10 (3.5%)	7 (0.9%)	0.01
Not Hispanic/Latino	1707 (77.1%)	811 (76.7%)	215 (74.9%)	596 (77.4%)	

	BLC2001 Screened n=2214	Bridging Eligible n=1057	Bridging Tested n=287	Not Tested n=770	p Value (Tested vs. Not Tested)
Not reported/Unknown	472 (21.3%)	229 (21.7%)	62 (21.6%)	167 (21.7%)	
Region, n (%)					
North America	350 (15.8%)	151 (14.3%)	50 (17.4%)	101 (13.1%)	0.20
Asia	346 (15.6%)	128 (12.1%)	33 (11.5%)	95 (12.3%)	
Europe	1518 (68.6%)	778 (73.6%)	204 (71.1%)	574 (74.5%)	
% Tumor area					
Mean (SD)	61.9 (30.37)	69.3 (23.66)	69.0 (24.20)	69.4 (23.47)	0.82
% Viable cells					
Mean (SD)	64.9 (25.18)	70.8 (22.48)	73.5 (20.85)	69.7 (22.98)	0.01
Site of tumor collected, n (%)					
Primary	1871 (84.5%)	935 (88.5%)	244 (85.0%)	691 (89.7%)	0.07
Metastatic	341 (15.4%)	121 (11.4%)	43 (15.0%)	78 (10.1%)	
Unknown	2 (0.1%)	1 (0.1%)	0 (0.0%)	1 (0.1%)	

D. Safety and Effectiveness Results

1. Safety Results

The safety of the *therascreen*® FGFR RGQ RT-PCR device is related to its accuracy as false results may lead to inappropriate treatment decisions. A false negative result would prevent a patient from receiving a potentially beneficial therapeutic. A false positive result would potentially expose the patient to a therapeutic that may not be beneficial as well as any possible side effects associated with the therapeutic.

The safety with respect to treatment with BALVERSA™ (erdafitinib) will not be addressed in detail in this SSED. Refer to USPI for safety information on BALVERSA™ (erdafitinib). The drug label also contains warnings and precautions about ocular disorders, hyperphosphatemia, and embryo fetal toxicity, associated with the use of the drug.

The accuracy of the FGFR kit relative to the comparator was determined using specimens from the 42756493BLC2001 study supplemented with procured specimens. The estimated PPA, NPA and OPA between the FGFR Kit and comparator (with comparator as the reference method) were 99.04%, 97.52% and 98.04% respectively, demonstrating that the FGFR Kit has high accuracy when compared to comparator and that the possibility of false results is very low.

Additionally, the use of the FGFR kit poses minimum safety hazard to patients, as biopsy specimens are routinely used in diagnosis and staging of UC patients.

2. Effectiveness Results – based on FGFR alteration detection with FGFR Kit

Analytical Concordance of the CTA and the FGFR Kit

As indicated in Figure 1 of Section B, specimens from 287 patients were eligible and tested by the FGFR kit in the bridging study.

For the 279/287 (97.2%) subjects that had valid results, the estimated PPA, NPA and OPA between the FGFR Kit and the CTA (with CTA as the reference method) are presented in Table 14 along with 95% CIs.

Table 14: Agreement between CTA and FGFR Kit (CTA as Reference Method)

	CTA		
FGFR Kit	**FGFR +**	**FGFR – ***	**Total**
FGFR +	69	6	75
FGFR –	12	192	204
Invalid	6	2	8
Total	87	200	287
PPA (95% CI) without Invalid		85.2% (75.9% - 91.3%)	
NPA (95% CI) without Invalid		97.0% (93.5% - 98.6%)	
OPA (95% CI) without Invalid		93.5% (90.0% - 96.1%)	
PPA (95% CI) with Invalid	-	79.3% (69.6% - 86.5%)	-
NPA (95% CI) with Invalid	-	96.0% (92.3% - 98.0%)	-
OPA (95% CI) with Invalid	-	90.9% (87.0% - 94.0%)	-

* Prior chemotherapy information was not collected for CTA negative patients; therefore 200 CTA negative subjects may include both Chemo Relapsed/Refractory and Chemo Naïve patients.

There were six subjects who were FGFR alteration positive by the FGFR Kit and were FGFR alteration negative by the CTA; three were p.G370C (c.1108G>T) positive, two were p.S249C (c.746C>G) positive and one was p.R248C (c.742C>T) positive with the FGFR Kit.

There were 12 subjects who were FGFR alteration negative by the FGFR Kit and were FGFR alteration positive by the CTA; one was FGFR3-BAIAP2L1 positive, two were FGFR2-BICC1 positive, two were FGFR2-CASP7 positive, one was p.S249C (c.746C>G) positive, one was p.G370C (c.1108G>T) positive, two were FGFR3-TACC3v1 positive, two were FGFR3:TACC1v3 positive and one was FGFR2-CASP7 plus FGFR3-TACC3v3 positive by the CTA. One subject (p.S249C (c.746C>G) positive) had a confirmed partial response to BALVERSA™ (erdafitinib); one was unevaluable, five had stable disease and five had progressive disease.

For 16/18 subjects with discordant results between the CTA and CDx, enough RNA was available for testing with a comparator method. In 13 of the 16 subjects: three CTA FGFR alteration negative and 10 CTA FGFR alteration positive, the comparator result agreed with the FGFR Kit result.

Clinical Concordance/Efficacy Analysis

ORR for CTA+ subjects

The ORR for the 87 Chemo Relapsed/Refractory Population was 32.2% (95% CI: 22.4%, 42.0%) by blinded independent review committee (BIRC).

ORR with FGFR Kit+ subjects

The primary objective of the clinical efficacy analysis was to estimate the ORR of subjects who were FGFR alteration positive with the FGFR Kit (FGFR Kit+). The observed ORR in treated subjects with a FGFR Kit+ result was estimated as a proportion of subjects with complete or partial response in all treated subjects with a FGFR Kit+ result and an FGFR alteration positive result with the CTA (CTA+). The ORR in all subjects who were positive by the FGFR Kit was estimated by weighting ORR over subjects with CTA+/FGFR Kit+ and subjects with CTA-/FGFR Kit+ results. Since prior therapy information for CTA negative patients was not collected in the clinical trial, NPA (97.0%) evaluated from 200 FGFR negative samples and CTA positive rate (21%) in the total screening population (Chemo Relapsed/Refractory and Chemo Naïve, 417/(417+1570) = 21%) were used to estimate the ORR. Based upon the biology of the disease, it is believed that response or resistance to chemotherapy will not impact the lack of benefit from erdafitinib in subjects with FGFR-negative (CTA negative) urothelial cancer. Therefore, the entire CTA-negative population could be used as the representative of the population of CTA negative chemo Relapsed/Refractory patients.

For subjects in the Chemo-Relapsed/Refractory Population (n=69), the ORR with FGFR Kit+/CTA+ alterations by BIRC was 33.3% (95% CI: 23.4% - 45.1%).

ORR with individual FGFR Kit+/CTA+ alterations

The ORR estimates by FGFR alterations (any point mutations, any fusion and any specific alterations) were also tabulated in Table 15. The specimens from six subjects harbored more than one alteration. These are summarized below:

- FGFR3-BAIAP2L1/p.S249C (c.746C>G)/FGFR3-TACC3v1 (1)
- p.G370C (c.1108G>T)/p.S249C (c.746C>G) (1)
- P.R248C (C.742C>T)/FGFR3-TACC3v1 (1)
- p.S249C (c.746C>G)/FGFR3-TACC3v1 (2)
- FGFR3-TACC3v1/FGFR3-TACC3v3 (1)

For the purposes of the ORR analysis, each alteration was included in the appropriate alteration totals separately, but only counted once in the overall point mutations or fusion total.

Table 15: ORR assessed by BIRC of Treated Subjects with FGFR+ by FGFR Kit and CTA Assays: Chemo-Relapsed/Refractory Population

	# Positive	# Response	ORR (95% C.I.)
FGFRi +	69	23	33.3% (23.4% - 45.1%)
Point Mutation +	58	22	37.9% (26.6% - 50.8%)
p.R248C (C.742C>T)	9	3	33.3% (12.1% - 64.6%)
p.S249C (c.746C>G)	37	13	35.1% (21.8% - 51.2%)
p.G370C (c.1108G>T)	3	1	33.3% (6.1% - 79.2%)
p.Y373C (c.1118A>G)	10	5	50% (23.7% - 76.3%)
Fusion+	15	2	13.3% (3.7% - 37.9%)
FGFR2-BICC1	0	0	-
FGFR2-CASP7	0	0	-
FGFR3-BAIAP2L1	1	0	0% (0% - 79.3%)
FGFR3-TACC3v1	11	2	18.2% (5.1% - 47.7%)
FGFR3-TACC3v3	5	0	0% (0% - 43.4%)

As there was no ORR data for FGFR2-BICC1, FGFR2-CASP7 and FGFR3-BAIAP2L1 fusions, it is not possible to claim that clinical validity for these alterations has been demonstrated.

ORR with FGFR Kit+ results

As the ORR in subjects with CTA-/FGFR Kit+ results is unknown because these subjects were not enrolled in 42756493BLC2001, a range of assumed ORR values of observed ORR in subjects with CTA+/FGFR Kit+ results were used to estimate a range of weighted overall ORR in subjects with FGFR Kit+ results. Bootstrapping (random sampling with replacement, 2000 times) was performed to calculate 95% CIs of the weighted overall ORR.

To estimate the weighted ORR in all FGFR Kit+ subjects, a range of ORR values for the positive discordant subjects, i.e., FGFR Kit+ but CTA-, were assumed as 100%, 75%, 50%, 25%, and 0% of the observed ORR in the positive concordant subjects, i.e., FGFR Kit+ and CTA+.

The weighted overall ORR in all FGFR Kit+ subjects in the Chemo-Relapsed/Refractory Population were 33.3% (95% CI: (21.7% - 45.1%), 32.3% (95% CI: 21.2% - 43.4%), 31.4% (95% CI: 20.7% - 42.1%), 30.4% (95% CI: 20.1% - 41.1%) and 29.4% (95% CI: 19.3% - 40.0%) when a range of ORR values for the positive discordant subjects (CDx+/CTA-), were assumed as 100%, 75%, 50%, 25%, and 0% of the observed ORR in the positive concordant subjects (CDx+/CTA+), shown in Table 16.

Table 16: ORR for FGFR+ by CDx Assays in Chemo Relapse/Refractory Subjects

	Assumed ORR of Patients with CDx+/CTA- Rate (%)	All Patients with CDx+ (obs, assumed) Rate (95% CI)
100% × Observed ORR in CDx+/CTA+	33.3%	33.3% (21.7% - 45.1%)
75% × Observed ORR in CDx+/CTA+	25.0%	32.3% (21.2% - 43.4%)
50% × Observed ORR in CDx+/CTA+	16.7%	31.4% (20.7% - 42.1%)
25% × Observed ORR in CDx+/CTA+	8.3%	30.4% (20.1% - 41.1%)
0% × Observed ORR in CDx+/CTA+	0.0%	29.4% (19.3% - 40.0%)

As conditions of approval, additional clinical validation data for FGFR2 fusions (FGFR2-BICC1 and FGFR2-CASP7) and FGFR3 fusion (FGFR3-BAIAP2L1) will be obtained from the ongoing phase 3 clinical trial.

3. Sensitivity Analysis

To assess the impact of missing CDx results on the estimate of the overall weighted ORR, the sensitivity analyses under 3 scenarios of concordance were performed. The best concordance case was to assume all missing CDx test results were concordant with the respective CTA results while the worst concordance case was to assign discordant results to all missing CDx values. In the likely concordance case, a logistic regression was built based on complete dataset as a classifier to predict missing CDx results. The covariates in building the classifier include CTA results, age, race, gender, region, percent tumor area, percent viable cells, and collection site of tumor. The fitted classifier yields a predicted probability of CDx positive result for each missing CDx sample. If predicted probability was greater than 0.5, the imputed CDx call for the sample was positive. Otherwise it was negative.

Eight of 287 subjects (2.79%) selected for the Bridging Study did not have a valid QIAGEN CDx result. Given the small number of missing values, minimal impact on the estimate of weighted overall ORR was expected. Insignificant changes in weighted overall ORR estimates were observed across the best-concordance, worst-concordance, and likely-concordance cases.

4. Pediatric Extrapolation

In this premarket application, existing clinical data was not leveraged to support approval of a pediatric patient population.

E. Financial Disclosure

The Financial Disclosure by Clinical Investigators regulation (21 CFR 54) requires applicants who submit a marketing application to include certain information concerning the compensation to, and financial interests and arrangement of, any clinical investigator conducting clinical studies covered by the regulation. The pivotal clinical

study included 87 investigators. None of the clinical investigators had disclosable financial interests/arrangements as defined in sections 54.2(a), (b), (c), and (f). The information provided does not raise any questions about the reliability of the data.

XI. SUMMARY OF SUPPLEMENTAL CLINICAL INFORMATION

None.

XII. PANEL MEETING RECOMMENDATION AND FDA'S POST-PANEL ACTION

In accordance with the provisions of section 515(c)(3) of the act as amended by the Safe Medical Devices Act of 1990, this PMA was not referred to the Molecular and Clinical Genetics Panel of Medical Devices, an FDA advisory committee, for review and recommendation because the information in the PMA substantially duplicates information previously reviewed by this panel.

XIII. CONCLUSIONS DRAWN FROM PRECLINICAL AND CLINICAL STUDIES

A. Effectiveness Conclusions

The primary endpoint for 42756493BLC2001 is ORR. Secondary efficacy endpoints include DOR, PFS, and OS. Response was assessed by investigators for all subjects. An independent radiologic review was also performed, given that the primary objective was achieved.

The primary efficacy population includes subjects who received at least one dose of study drug in the trial and had relapsed or were refractory after at least 1 prior line of chemotherapy (n=87). The ORRs for treated subjects in the Chemo-Relapsed/Refractory Population who were CTA+ or FGFR Kit+, as assessed per BIRC, are presented in Table 17. In all analyses, the time to response was short; at approximately 1.4 months (i.e. the first disease evaluation time point), while other time-to-event outcomes such as median DOR (approximately 5.5 months), median PFS (approximately 5.6 months) and median OS (approximately 12 to 14 months) were consistent, indicating that the efficacy outcomes overall were robust.

Table 17: ORR for CTA+ vs FGFR Kit+ Treated Chemo-Relapsed/Refractory Subjects

	Total Subjects (n)	ORR (95% CI) BIRC
CTA+	87	32.2% (22.4 – 42.0)
FGFR Kit+/CTA+	69	33.3% (23.4 – 45.1)
FGFR Kit+*	-	33.3% (23.4 – 45.1)
FGFR Kit+&	-	29.4% (19.3% - 40.0%)

* FGFR Kit+ but CTA- subjects, were assumed as 100% of the observed ORR in the FGFR Kit+ and CTA+ subjects;

[&] FGFR Kit+ but CTA- subjects, were assumed as 0% of the observed ORR in the FGFR Kit+ and CTA+ subjects

Moreover, significant responses were observed regardless of pre-treatment status, as responses were observed in patients with none, one, or two prior lines of therapy. Responses were also observed in patients with visceral metastases and in patients who had received prior PD-L1/PD-1 directed therapy.

B. Safety Conclusions

The FGFR Kit is not expected to directly cause actual or potential adverse effects, but test results directly impact patient treatment risks. The risks of the FGFR Kit are associated with the potential mismanagement of patients resulting from false results of the test. Failure of the device to perform as expected or failure to correctly interpret test results may lead to incorrect FGFR test results, and consequently improper patient management decisions in UC treatment. A patient with a false positive result may undergo treatment with BALVERSA™ (erdafitinib) with inappropriate expectation of therapeutic benefit. A patient with a false negative result may be treated without BALVERSA™ (erdafitinib) and not experience the potential benefit. Analytical performance in this submission demonstrates that the assay is expected to perform with high accuracy mitigating the potential for false results.

C. Benefit-Risk Determination

The probable benefits of the FGFR Kit are based on data collected in the clinical Bridging Study conducted to the 42756493BLC2001 clinical trial of BALVERSA™ (erdafitinib). The FGFR Kit demonstrated overall agreement to the CTA of 93.6% (90.0% - 96.1%). The ORR for the Primary Efficacy Population who were FGFR Kit+ was 33.3% (95%CI: 23.4, 45.1) by BIRC assessment.

The risks of the FGFR Kit are associated with the potential mismanagement of patients resulting from false results of the test or a failure to receive results. Patients who are determined to be false positive by the test may be exposed to a drug that is not beneficial and has adverse events. A false negative result may prevent a patient access to a potentially beneficial drug. The likelihood of false results was assessed in the analytical evaluation and showed acceptable performance with overall agreement to a comparator method 97.72% (worst concordance case) when considering test invalids and 98.04% when excluding test invalids. The reported PPA and NPA was 99.04% and 97.04%, respectively including test invalids (worst concordance case). Excluding test invalids, the PPA and NPA were 99.04% and 97.52%, respectively.

Treatment with BALVERSA™ (erdafitinib) provides meaningful clinical benefit measured by overall response rate, with an acceptable safety profile.

In conclusion, given the available information above, the data supports that in selecting specific FGFR alteration positive UC patients using the FGFR Kit for treatment with BALVERSA™ (erdafitinib), the probable benefits outweigh the probable risks.

1. Patient Perspectives

This submission did not include specific information on patient perspectives for this device.

D. Overall Conclusions

Study 42756493BLC2001 met its primary objective, demonstrating a clinically meaningful ORR in subjects with advanced UC who have certain FGFR genetic alterations. Responses to BALVERSA™ (erdafitinib) treatment were rapid and durable and were independent of the number of prior lines of therapy, presence of visceral metastases, or tumor location.

BALVERSA™ (erdafitinib), with personalized dose modifications, was well tolerated as a single agent. The safety profile shows that primarily FGFRi-specific, non-systemic toxicities were generally reversible, usually not dose limiting, and could be managed by supportive care and dose modification.

Collectively, the efficacy and safety results from the 42756493BLC2001 study demonstrate that BALVERSA™ (erdafitinib) has a favorable benefit-risk profile and is a useful treatment in subjects with advanced UC who have certain FGFR genetic alterations.

The data in this application support the reasonable assurance of safety and effectiveness of this device when used in accordance with the indications for use. Data from study 42756493BLC2001 support the validity of FGFR Kit as an aid in selecting patients with UC with specific FGFR alterations for whom BALVERSA™ (erdafitinib) is indicated.

XIV. CDRH DECISION

CDRH issued an approval order with conditions on April, 12, 2019.
The applicant's manufacturing facilities have been inspected and found to be in compliance with the device Quality System (QS) regulation (21 CFR 820).

XV. APPROVAL SPECIFICATIONS

Directions for use: See device labeling.

Hazards to Health from Use of the Device: See Indications, Contraindications, Warnings, Precautions, and Adverse Events in the device labeling.

Post-approval Requirements and Restrictions: See approval order.

XVI. REFERENCES

1. Helsten T, et al., 2015. The FGFR Landscape in Cancer: Analysis of 4,853 Tumors by Next-Generation Sequencing. Clin Cancer Res, 22(1), pp.259-67
2. Rodriguez-Vida A, *et al.*, 2015. Complexity of FGFR Signalling in Metastatic Urothelial Cancer. *J Hemetol Oncol*, 8, pp.119
3. Knowles, MA & Hurst, CD, 2014. Molecular Biology of Bladder Cancer: New Insights into Pathogenesis and Clinical Diversity. *Nat Rev Cancer*, 15(1), pp.25-41