Improved tumor discrimination and shortened administration-to-imaging times in fluorescence guided surgery through paired-agent protocols

Cheng Wang, Margaret Folaron, Jason R. Gunn, Sassan Hodge, Eunice Y. Chen, P. Jack Hoopes, Kenneth M. Tichauer, Kimberley S. Samkoe

1. Thayer School of Engineering, Dartmouth College, Hanover, NH
2. Department of Surgery, Dartmouth-Hitchcock Medical Center, Lebanon, NH
3. Geisel School of Medicine, Dartmouth College, Hanover, NH
4. Biomedical Engineering, Illinois Institute of Technology, Chicago, IL

*Corresponding Author: Kimberley S. Samkoe, Kimberley.S.Samkoe@dartmouth.edu, phone: 603-650-7618
Abstract

Background: The goal of fluorescence guided surgery (FGS) in oncology is to improve the surgical therapeutic index by enhancing contrast between cancerous and healthy tissue. However, optimal discrimination between these tissues is complicated by the non-specific uptake and retention of molecular targeted agents and the heterogeneity of fluorescence signal. Paired-agent imaging (PAI) employs co-administration of an untargeted imaging agent with a molecular targeted agent, providing a normalization factor to minimize nonspecific and heterogeneous signals. The resulting measured binding potential is quantitative and equivalent to in vivo immunohistochemistry of the target protein. This study demonstrates that PAI improves the accuracy of tumor-to-healthy tissue discrimination compared to single agent imaging for in vivo FGS.

Methods: PAI using a fluorescent anti-EGFR affibody molecule (ABY-029, eIND 122681) with untargeted IRDye 700DX carboxylate was compared to ABY-029 alone in an oral squamous cell carcinoma xenograft mouse model at 3 hours (n = 30).

Results: PAI significantly enhanced tumor discrimination, as compared to ABY-029 alone (diagnostic accuracy - 0.94 vs 0.86, ROC curve AUC - 0.991 vs. 0.925, respectively). Additionally, the AUC of the ROC curve for PAI was stable as patient cohort number was increased from n = 1 to 20, while ABY-029 decreased as n increased, indicating a potential for universal FGS image thresholds to determine surgical margins. In addition, PAI reduced the administration-to-imaging time from 3 hours to 30 minutes and exhibited a statistically stronger correlation to EGFR expression heterogeneity ($r = 0.58$ compared to $r = 0.47$, $p = 9 \times 10^{-8}$).

Conclusion: The quantitative receptor delineation of PAI promises to improve the surgical therapeutic index of cancer resection in a clinically relevant timeline.
Keywords: Paired-agent imaging, fluorescence guided surgery, Head and neck squamous cell carcinoma, epidermal growth factor receptor, ABY-029, IRDye 700DX
**Background**

Completeness of surgical resection is a critical determinant for the survival of patients with head and neck cancers. Residual tumor increases tumor-related death at 5-years by 90% compared to those with truly negative margins (1) but the use of wide margins to remove residual tumor can lead to severe morbidity. The near ubiquitous overexpression of epidermal growth factor receptor (EGFR) - with estimates of > 90% overexpression in SCC (2,3) - has led to the development of numerous molecular therapeutic agents, which have been subsequently leveraged for fluorescence imaging (4–8). The goal of molecular-targeted fluorescence-guided surgery (FGS) is to improve the surgical therapeutic index based on the overexpression of the molecular target in tumor compared to normal tissue. Several studies have reported advantages of FGS for identifying bulk tumor and tumor margins using therapeutic antibodies (cetuximab, panitumumab) labeled with IRDye 800CW (LI-COR Biosciences, Inc.) (7,9–11). However, true molecular contrast using FGS is confounded by heterogeneous uptake and nonspecific retention of targeted imaging agents within all tissue types. Paired-agent imaging (PAI) methods have the potential to overcome these confounding effects through co-administration of a second, untargeted, control fluorescent agent enabling imaging of the receptor concentration, rather than agent concentration. This pre-clinical project compares the accuracy of tumor discrimination using conventional “single-agent imaging” (SAI) and the proposed PAI in an orthotopic xenograft mouse model of human head and neck cancer.

In recent years, FGS has been improved by two advancements: the aforementioned PAI and the development of an anti-EGFR fluorescent Affibody molecule (ABY-029). PAI, which reports the “binding potential” (BP, a value proportional to receptor concentration), has been used in a variety of EGFR-overexpressing xenograft cell lines to demonstrate that tumor-averaged binding potential scales linearly with EGFR both in vivo and ex vivo (12). This in vivo phenomena was
linearly correlated with \textit{ex vivo} tumor EGFR immunohistochemistry (13) and shown to non-invasively detect fewer than 200 tumor cells in draining lymph nodes (14). However, the ability of PAI to truly improve tumor discrimination in FGS has never been quantitatively assessed.

ABY-029 was developed to decrease administration-to-imaging time (hours instead of days) and reduce immunogenicity compared to antibody imaging agents (6). ABY-029 is currently being tested in Phase 0 studies in a number of solid tumor types, including head and neck cancers. In the work presented here, we utilize two orthotopic SCC base of tongue tumors (FaDu and Detroit 562) and a highly expressing EGFR SCC of the skin (A431) to compare the accuracy and efficiency of FGS tumor resection using ABY-029 alone versus PAI with ABY-029 in combination with control IRDye 700DX carboxylate.

\textbf{Methods}

\textit{Cell lines and culture methods}

Human squamous cell carcinoma cell lines used in this study included FaDu, a pharynx carcinoma, Detroit 562, a metastatic pharynx carcinoma derived from pleural effusion, and A431, an epidermal SCC. All three cell lines were purchased from the ATCC (Manassas, VA, USA) and were cultured according to ATCC specifications with the addition of 1\% penicillin-streptomycin.

\textit{Imaging agents}

ABY-029 was obtained from the University of Alabama at Birmingham (UAB) Vector Production Facility and manufactured under Good Laboratory Practice (GLP) as previously described (6). The ABY-029 human microdose is defined as 30 nanomole per human, (3.96 \( \mu \)g/kg for a 60 kg human).
Using the method of Reagan-Shaw (2007)(15), the mouse-equivalent dose was determined to be 48.8 µg/kg for an average 22 g mouse, for a final dose of 1.07 µg/mouse. IRDye 700DX NHS ester was purchased from LI-COR Biosciences, Inc. (Lincoln, NE) and converted to carboxylate form by dissolving in PBS (pH = 8.5) and stirring at room temperature for 5 hours.

**Mouse xenograft model**

All animal procedures were approved by the Dartmouth Institutional Animal Care and Use Committee (IACUC) and conducted according to NIH-OLAW and AAAALAC guidelines. Female, athymic nude mice, 6-8 weeks of age, were purchased from Charles River Laboratories (Wilmington, MA). Tongue tumors were implanted using a 25-gauge needle and 5×10^5 cells in 50 µL culture medium. Ten mice were implanted per cell line (n_total = 30). PAI was performed on six mice for FaDu, and seven each for Det 562 and A431 tumor lines, and the remaining mice were used for auto-fluorescence imaging. An additional 30 mice were implanted with FaDu for the administration-to-imaging time study (n = 5 mice/time point). Tumor take was 100% and were used for imaging when the diameter reached ~3-4 mm.

**ABY-029 and IRDye 700DX fluorescence imaging**

Mice were administered 200 µL of a 1:10 molar ratio of 0.68 µM ABY-029 and 6.8 µM IRDye 700DX in sterile phosphate buffered saline (PBS) via intravascular tail vein injection (Fig 1A). Injection concentrations were determined by assessing fluorescence signal at three hours compared to non-injected autofluorescence of the tissue (Supplemental Data, Section S1). The mice were sacrificed by cervical dislocation while anesthetized to a surgical plane (1.5-2% isoflurane, 1
L/min O2) 3-hours post-administration, with the exception of the time study where sacrifice occurred at 0.25, 0.5, 1, 2, and 5 hours. Note that the 3-hour FaDu tumors were also used in the time study to minimize animal use. The tongue was excised at the base, bisected along the raphe, and placed on a glass slide cut-face down. A section of normal leg muscle (EGFR-negative control) was also included. *Ex vivo* images of the tissues were collected for both ABY-029 and IRDye 700DX carboxylate on the Odyssey CLx (LI-COR Biosciences, Inc.) using the following settings: auto function for laser intensity, 1-mm focus offset, medium quality, and 42-μm resolution.

**PAI binding potential map creation**

PAI binding potential (BP) maps were calculated from of the ABY-029 and IRDye 700DX images using the single-time point (STP) method (Equation 1), first described by Tichauer et al (16) (Fig 1B). Note that in this previous work, pre-injection images were subtracted from the post-injection image to remove the contribution of autofluorescence. Pre-administration images were not subtracted and autofluorescence contribution is discussed in the *Supplemental Data, Section S1*. For each pixel within the image, the BP was calculated using

\[
BP = \frac{I_T}{I_U \times NF} - 1
\]  

(1)

where \(I_T\) and \(I_U\) are the pixel intensity of the targeted (ABY-029) and untargeted (IRDye 700DX) imaging agents, respectively, and \(NF\) is the normalization factor determined by Equation 2.

\[
NF = (BP_{\text{norm}} + 1) \frac{\bar{I}_T(norm)}{\bar{I}_U(norm)} = \frac{1.5 \bar{I}_T(norm)}{\bar{I}_U(norm)}
\]  

(2)

where \(\bar{I}_T(norm)\) and \(\bar{I}_U(norm)\) are the mean pixel intensities of the targeted and untargeted imaging agents, respectively, in the normal tongue. The \(NF\) is calculated independently for every ABY-029 and IRDye 700DX image pair. The \(BP_{\text{norm}}\) is the binding potential of the normal tongue, which was artificially set to 0.5 in order to have a “near-zero” BP value in the normal tongue tissue while
avoiding negative pixels. The selection of the NF is described in detail in the Supplemental Data, Section S2, where it is demonstrated that the selection of the NF does not alter the observed contrast.

Pathology

After imaging, the tongue sections were placed on filter paper to maintain orientation and fixed in 10% buffered formalin (Biochemical Science, Inc) in histological cassettes. Standard H&E and EGFR immunohistochemistry (IHC) staining were performed by the Norris Cotton Cancer Center Pathology Translational Research Resource as described previously (13). RGB images of whole H&E and EGFR IHC tissue sections were collected on the Vectra 3 (Perkin Elmer) at 4X magnification. The stitched overview image was saved as an RGB three image stack .tiff file, then converted to a single .tiff file using the concatenate arrays function (cat) in MATLAB version R_2017a.

Image Preparation and Co-registration

Five image types are used for in this study: H&E, EGFR IHC, ABY-029, IRDye 700DX, and BP maps. Prior to alignment and co-registration, several steps were taken to prepare the images. ABY-029, IRDye 700DX and BP maps were inherently co-registered as a function of imaging on the Odyssey CLx (42 μm/pixel). The H&E and EGFR IHC images collected on the Vectra scanner (1 μm/pixel) were resized to match the fluorescent images. The brown stain indicating EGFR in the IHC images was isolated by using the H DAB Color Deconvolution script in FIJI (17) and then normalized to the average stain intensity in the placenta positive control slide for each staining batch to correct for variations in stain intensity due to color development. Image alignment and co-registration of the fresh tissue sections with pathology was performed using a previously described procedure (18). Briefly, the BP map was co-registered to the EGFR IHC image using
*warp_it* in MATLAB that utilizes point set registration and point matching to spatially transform and align the images. Visualization of the overlaid images is provided in the *Supplemental Data, Section S3*. The H&E were not used for image analysis, and thus the coordinate transformation determined for the IHC was applied to the H&E image for visualization purposes only.

**Image Analysis and Statistics**

For each sample, a pathologist drew regions-of-interest (ROI) for normal tongue muscle, tumor, and salivary gland using EGFR IHC, which were manually converted and translated to co-registered fluorescence and BP images for both mean and pixel-by-pixel analysis of tissue type. The visualization of ABY-029, IRDye 700, and BP are presented by “fire,” “kryptonite,” and “ice” pseudo colormaps as defined by *COLORMAP* ([https://jdherman.github.io/colormap/](https://jdherman.github.io/colormap/)). Histograms were created in OriginPro 2018 (OriginLab). Receiver operating characteristic (ROC) curves were calculated with either *percurve* function in Matlab, or ROC analysis in OriginPro. Statistically significant differences in group means were analyzed in OriginPro with a one-way ANOVA with Bonferroni correction to avoid Type I error. Diagnostic parameters, including area under the-curve (AUC), sensitivity, specificity, positive prediction value (PPV), negative prediction value (NPV), and diagnostic accuracy were determined. To study the correlation of EGFR heterogeneity, pixel intensity values were extracted from deconvoluted IHC images, fluorescent images, and BP maps and were normalized to the highest pixel value. To study the impact of image resolution on the correlation of EGFR staining and fluorescent images, *imagepyramid* in Matlab was used for averaging four adjacent pixel values. Pixel-by-pixel correlation was performed using Pearson product correlation \(r\) in OriginPro. Contrast-to-Variance Ratio (CVR) was defined by:

\[
CVR = \frac{\mu(I_T) - \mu(I_N)}{\sqrt{\sigma_T^2 + \sigma_N^2}}
\]  

(3)
\( \mu(I_T) \) and \( \mu(I_N) \) represent mean fluorescence or BP, and \( \sigma_T \) and \( \sigma_N \) represent the standard deviation of fluorescence or BP values in tumor and normal tissue, respectively. Statistical significance of BP and fluorescence intensity over time was calculated using one-way repeated measures ANOVA.

**Results**

*Administration dose and the normalization factor*

In previous studies (13,14), a pre-injection background image was used to remove the absolute tissue autofluorescence signal on a pixel-by-pixel basis; however, in this study and during FGS process, pre-injection images within the excised tissue and surgical wound bed are not available. This had two effects on study design. First, autofluorescence was used to determine the appropriate administered paired-agent dose. A 1:10 molar ratio of ABY-029:IRDye 700DX was used in order to obtain fluorescence signal \( \sim \)8 times the autofluorescence at 3 hours (see *Supplemental Figure S1 & S2*). Second, the normalization factor (NF) was determined at each time point by setting the BP equal to 0.5 by using the average ABY-029 and IRDye 700DX fluorescence signal in normal tongue, rather than using EGFR-devoid leg muscle in the first post-administration image. CVR (Equation 3) was used to standardize the measurements and compare image contrast between PAI and SAI as tumor-to-background ratios (TBR) were found to be unstable (see *Supplemental Figure S3*).

**Discrimination of tumor and normal tissue**

To evaluate the ability of PAI and SAI to distinguish tissue types based on signal alone, we analyzed the resultant images in two ways: region of interest (ROI) averages and region pixel-to-pixel
comparisons. After co-registration, the fluorescence intensity of ABY-029 and the BP were compared for tumor (FaDu, Detroit 562 and A431), normal tongue, salivary gland if present, and leg muscle (negative control, Fig. 2A). A representative example of ABY-029 targeted fluorescence intensity and BP is plotted in Fig. 2B to demonstrate the variability of the signal in each tissue. The ROI-specific averages for each tissue type are shown in increasing order of EGFR expression (Fig. 2C). A one-way ANOVA analysis with Bonferroni correction demonstrated that all fluorescence means were not the same for ABY-029 fluorescence ($p < 9 \times 10^{-26}$) or BP ($p < 2 \times 10^{-23}$). The Bonferroni individual means comparison showed the ABY-029 fluorescence of Detroit 562 and A431 tumors were significantly different from all normal tissues, while FaDu, salivary gland, and normal tongue were not significantly different from each other. Leg muscle, which lacks EGFR expression, has significantly lower ABY-029 fluorescence compared to all other tissues. Individual means comparison for BP showed that all tumor lines were significantly different from all normal tissues, all normal tissues were significantly different from each other, but FaDu and Detroit 562 could not be statistically distinguished.

Although tumor and normal tissues could be identified based on BP means, the histograms in Fig. 2B demonstrate a substantial overlap in the distribution of pixel values between some groups. Therefore, tissue type comparison of SAI and PAI was performed on the co-registered images on a pixel-to-pixel basis (Fig. 3). Representative co-registered IRDye 700DX fluorescence, ABY-029 fluorescence, and BP map images are shown for each tumor line with pathological images (Fig. 3A). Co-registered pixel intensities from the EGFR IHC (gold standard) and fluorescence images were used to plot receiver-operating characteristic (ROC) curves to evaluate the diagnostic ability of both SAI and PAI. For each tumor type, BP maps yielded higher area-under-the curve (AUC) values than either ABY-029 and IRDye 700DX alone. When mice were grouped into
cohorts of individual tumor type and all tumor lines (Fig. 3B), the BP AUCs were ≥ 0.990, while ABY-029 AUCs were consistently lower. The cohort diagnostic accuracy statistics are summarized in Fig. 3B, with the higher value highlighted in green for ease of interpretation. BP maps demonstrated statistically higher ($p < 0.05$) diagnostic ability in all cases, with the exception of Detroit 562, in terms of sensitivity and negative predictive value (NPV).

**Representation of tissue heterogeneity**

EGFR expression within tumors was highly heterogeneous, especially compared to normal tissues, as can be observed in the IHC images (Figs. 2 and 3). Heterogeneous EGFR expression can contribute to difficulties in distinguishing tissues; therefore, we assessed the pixel-by-pixel linear correlation between IHC stain intensity with BP, and ABY-029 and IRDye 700DX fluorescence (Fig. 4). To assess the effects of co-registration error on the high resolution (42 µm/pixel) images, an image pyramid algorithm (Fig. 4A) was used to incrementally decrease resolution. The resulting scatter plots, and the corresponding Pearson coefficients ($r$) for each resolution tested in a representative FaDu tumor are shown in Fig. 4A. At Level 0 (42 µm/pixel), all three image types exhibited a moderate correlation with IHC with BP demonstrating the strongest correlation while the non-targeted IRDye 700DX is negatively correlated. As image resolution is decreased, the strength of the correlation between EGFR IHC and all three image types non-targeted IRDye 700DX is negatively correlated increase, with BP maintaining the strongest correlation at each level. The correlation between the IHC and IRDye 700DX was negative and remained low (0.1-0.3) to moderate (0.3-0.5); therefore, only the correlations between IHC and BP or targeted fluorescence were further analyzed. The rate of change in $r$ is plotted in Fig. 4B and plateaued after a 1/8 reduction in resolution. Clinical imaging systems (wide-field to endoscopic) have spatial resolution of 50-
500 µm (grey shaded region, Fig. 4B). Therefore, a 1/4 reduction in the original resolution (168 µm), was applied in further analyses of each tumor group. Overall, the Pearson coefficients were 0.58 ± 0.04 and 0.47 ± 0.05 for BP and ABY-029, respectively (p < 0.0001). In all cases, IHC demonstrated a stronger correlation with BP than ABY-029 fluorescence alone (p < 0.005).

Reduction of administration-to-imaging time

To maximize observed FGS contrast, the delay time between agent administration and surgery must be optimized. To study administration-to-imaging time of PAI, mice (n = 5 or 6 per group) were co-administered ABY-029 and IRDye 700DX, then sacrificed at varying time points up to 5 hours after administration (Fig. 5). Representative SAI and PAI images for a single animal at each time point, as well as plots of the average signal intensity in the tumor region over all times are shown (Fig. 5A). The ability to distinguish the tumor region improved with time after injection for both SAI and PAI. The PAI BP yielded a stable CVR of 2.0 ± 0.2 by 30 min post-agent-injection, while ABY-029 alone required 3 h to reach a CVR plateau of only 2.2 ± 0.3 (Fig. 5B). PAI BP demonstrated a higher CVR compared to ABY-029 SAI at all time points (Fig. 5B); however, they are only significantly different during the first 2 hours determined by One-Way repeated measures ANOVA with time as a within-subjects variable (p < 0.05). ROC curves for BP and ABY-029 were plotted for each time point (Fig. 5C). The AUCs of both BP and ABY-029 alone increase over time; however, at each time point the PAI BP AUCs are higher than SAI AUCs. Moreover, SAI AUC increases over the 5-hour administration-to-imaging time with no demonstration of stabilization.

Discussion
Molecular PAI protocols have been proven to provide significant advantages for estimating true molecular contrast and for enabling unmatched specificity and sensitivity (13,14). As previously demonstrated (16,20), the tissue-averaged PAI binding potential (BP) in this work scaled with EGFR expression in tumor lines, and the means of all tumor lines studied were significantly higher compared to all normal tissues (Fig 2). In comparison, the mean fluorescence intensity from SAI with ABY-029 alone failed to be able to separate the lowest EGFR expressing tumor line from normal oral tissues. However, the broad heterogeneity of both BP and fluorescence intensity in PAI and SAI, respectively, warranted further investigation of the diagnostic abilities of these methodologies (Fig 2B). Therefore, studies mimicking in-patient and back-table intraoperative assessment strategies (2) were undertaken to compare PAI and SAI to gold standard H&E and EGFR immune-stained formalin-fixed paraffin embedded pathological specimens, with co-registration to PAI and SAI and correlation on a pixel-to-pixel basis (Fig 3).

In every category tested, including ROC curve determined AUC and diagnostic tests (Fig 3), PAI outperformed SAI with the exception of sensitivity and negative predictive value (NPV) for the Detroit 562 cell line (Fig 3B). Both sensitivity and NPV include “False Negatives” in the denominator, suggesting that Detroit 562 had a high number of pixels within the pathologist-designated tumor regions that were classified as normal tissue based on PAI as compared to SAI. This discrepancy can be explained by the fact that the “pathologist-determined tumor” contains regions of low-EGFR expressing tissue (Fig. 2B), and PAI is designed to enhance contrast as a function of targeted molecule (in this case, EGFR expression). When considering the whole tumor on a pixel-to-pixel basis, the regions devoid of EGFR decreased the measured predictive power of PAI because it is truly a molecular signal, unlike SAI that is a mixture of molecular targeted and non-specific signal from the enhanced permeability and retention (EPR) effect. In cases where EGFR
expression is low or there are large non-EGFR staining regions within the tumor, SAI may be able
to better differentiate tumor from normal tissue based on EPR effect. Data in soft-tissue sarcomas
that suggests this is the case by demonstrating that overall tumor contrast was enhanced and fluo-
rescent signal heterogeneity minimized by simultaneously imaging perfusion-based ICG accumu-
ation with ABY-029 in a single imaging channel (21,22). However, perfusion agents may be better
for this capacity. Further investigation is required in tumors with large negative regions, or with
lower cellular density—often seen in aggressive cell lines commonly used for xenograft models
(due to fast growth rate), yet not necessarily indicative of patient population tumor characteristics.

It is interesting to note, as we moved from ROC analysis of a single animal ($n = 1$ per cell
line, Fig. 3A) to a tumor cohort ($n = 6$ or 7 per cell line, Fig 3B) and finally, to all tumor cohort ($n
= 20$, Fig 3B) that the AUC of BP remained relatively constant (AUC $\geq 0.99$) while the AUC of
ABY-029 fluorescence decreased as the group size increased. This is important when considering
broad implementation of FGS into surgical suites. Fluorescence intensity alone can vary widely
patient-to-patient, owing to variability in fluorescent agent administration, delivery, and excretion,
which increased variability causing the sensitivity and specificity to decrease as the sample size
increases. The stability of PAI AUC of the ROC over increasing patient populations was likely
due to PAI ratio-metric imaging methodologies removing the hemodynamic variation of delivery
and clearance rates of the dye between individuals. BP calculated using PAI could be a promising
standard threshold for tumor region detection, a hypothesis that will be explored in future planned
clinical studies.

Tumor spatial heterogeneity is an important prognostic factor, and accurately imaging re-
ceptor expression heterogeneity is key for identifying tumor regions. This is especially important
when attempting to identify tumor in the surgical margins where cell density, and therefore EGFR
concentrations, may be low. As anticipated from previous studies, PAI and ABY-029 were positively correlated, while untargeted IRDye 700DX was negatively correlated with EGFR IHC (Fig 4B) (21). At 42-μm resolution, ABY-029 and PAI BP were only moderately correlated with gold standard IHC images (Fig 4B). This may be due to several factors. First, it can be observed that there is a large population of pixels in the IHC images that were clustered at the highest measurable pixel values (Fig 4B) due to the limited dynamic range of IHC images (0–2 OD). All three tumor lines had intense IHC staining and have been found to have moderate to high EGFR expression (1.2 (± 0.3) ×10^6, 1.6 (± 0.6) ×10^5, and 7.4 (± 0.4) ×10^4 EGFR receptors/cell in A431, Detroit 562, and FaDu, respectively; see Supplemental Table S1 & Fig S5) and received a pathologist score of 3+ with strong, continuous membranous staining. Secondly, the imperfect registration between fresh tissue fluorescence images and fixed pathological tissue images could substantially reduce correlation, especially at high image resolutions where single pixel misalignments are likely. This was validated when lower, but still clinically viable, image resolutions were examined using a 4×4 Gaussian filter and improvements in correlations for both PAI and ABY-029 SAI were observed (Fig 4). This indicates that better methodologies for spatial correlation of fresh tissue to pathologically stained sections are needed for single cell assessment of surgical margins.

Image contrast-to-variance between the tumor and the normal tissue depends on many factors, including the administration-to-image timing (to allow normal tissue clearance), the instrumentation used, the dose of fluorophore given, the health of the patient (e.g., diseased liver/kidney may extend plasma half-life), the physiology of the tumor and healthy tissue (e.g., blood flow and vascular permeability), the on- and off- rate constants of specific agent binding, the level of non-specific agent binding, and volume of tissue interrogated. In situ imaging, where decision making may be most critical, tends to exhibit lower image contrast-to-variance than excised tissues (23),
likely attributable to the nonspecific signal arising from the bulk normal tissue. There have been many strategies tested to increase contrast by decreasing non-specific signal. Administration of a pre-dose, or “cold dose”, of a non-fluorescent antibody an hour prior to a fluorescent antibody has been used to decrease contribution of molecular target sinks within the body. However, the results for these studies have been mixed. Moore et al.(24) and Voskuil et al. (7) demonstrated improved tumor-to-background ratios by pre-dosing whereas Nishio et al. (25) concluded that there was no additional value added.

A large body of work, including clinical trials, has been produced using high-dose ICG and “second window” administration-to-imaging times (AIT) i.e., 24 h AIT. The high dose of ICG allows sufficient tumor accumulation such that the tumor is visible at 24 hours, even with the fast plasma and normal tissue clearance typically observed. In addition, several groups have demonstrated the use of IRDye 800CW labeled EGFR-targeted antibodies with surgery and imaging at 1-4 days post-administration (4,5,24,25), with optimum fluorescence intensities observed within the first 2 days (25). On the other hand, ABY-029 (~8 kDa)—with its considerably faster plasma clearance half-life (~ 20 min) (6) as compared to monoclonal antibodies (~114 hours)—yields optimal AIT of only a few hours (21,27) as demonstrated here. However, the ability of PAI to provide stable contrast and high tumor differentiation starting at 30 min (and extending for hours) highlights the potential for in-surgical suite administration of PAI, reducing complexities in patient appointments and surgical timing that can occur with SAI.

Conclusions
PAI has the potential to broadly impact the clinical implementation of fluorescence guided surgery. PAI diagnosed tongue SCCs with high accuracy, which was less sensitive to inter-patient variability. PAI accurately represented the true molecular heterogeneity of receptor expression in tumors and could be used to image over a wide range of clinically applicable resolutions. In addition, PAI demonstrated the potential to facilitate flexibility within the surgical setting by decreasing the time from imaging agent administration to the start of resection while maintaining high diagnostic accuracy. We propose the use of PAI as an innovative molecular imaging method that will improve the diagnostic accuracy and efficiency of FGS.

List of abbreviations

FGS: Fluorescence guided surgery; PAI: Paired-agent imaging; EGFR: Epidermal growth factor receptor; SAI: Single-agent imaging; BP: Binding potential; GLP: Good Laboratory Practice; PBS: phosphate buffered saline; STP: Single-time point; NF: Normalization factor; ROI: Regions-of-interest; AUC: Under the curve; ROC: Receiver operating characteristic; CVR: Contrast-to-Variance Ratio; PPV: Positive prediction value; NPV: Negative prediction value; TBR: Tumor-to-background ratio; AIT: Administration-to-imaging time
Declarations

Ethics approval and consent to participate

All animal procedures were approved by the Dartmouth Institutional Animal Care and Use Committee (IACUC) and conducted according to NIH-OLAW and AAALAC guidelines.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Competing interests

KSS reports an ongoing academic-industrial relationship for the development and clinical testing of ABY-029 with LI-COR Biosciences and Affibody AB, and personal fees from LI-COR and American Institute for Biological Sciences.

Funding

This work was funded by the grant R37 CA212187 (KSS). The production of ABY-029 was previously funded by R01 CA167413, flow cytometry was performed with support from 5P30 CA023108-41.

Authors’ contribution

CW designed the research, performed experiments, analyzed data and wrote the manuscript. MF, JRG, SH, and PJH helped perform experiments, analyzed data, and/or interpreted data. EYC, KMT helped design the study and interpreted data. KSS designed the research, provided the funding, interpreted data, and supervised the experiments. All authors read, edited, and approved the final manuscript.

Acknowledgements
The authors would like to acknowledge the gift of ABY-029 (R01 CA167413, PI Keith Paulsen), EGFR quantification experiments were carried out in DartLab, the Immune Monitoring and Flow Cytometry Shared Resource at the Norris Cotton Cancer Center at Dartmouth, with NCI Cancer Center Support Grant with support from 5P30 CA023108-41.

REFERENCES


Figure Captions

Figure 1. Schematic of experimental procedures. (A) Mice, with xenograft murine tongue tumors, were administrated ABY-029 and IRDye 700DX by tail vain injection. Inset - a pictorial representation of the paired-agent distribution 3-hours after administration, where both agents are present due to non-specific binding and uptake in all tissues, while only ABY-029 specifically bound to receptors. After sacrifice, the tongue is removed and bisected. Tumor and normal tissue fluorescence was imaged using Odyssey CLx. (B) Binding potential (BP) maps representing available EGFR concentrations were created by ratioing targeted and normalized untargeted fluorescence. For further impartial analysis, pathological and fluorescence images were digitally aligned and co-registered.

Figure 2. PAI exhibits ability to distinguish normal and tumor tissues using average ROI signal intensity. (A) After image co-registration, pathologist-defined ROIs of tumor, normal tongue (NT), and salivary glands (SG) on H&E sections were translated manually to the digitally aligned images of EGFR IHC, ABY-029, and BP. (B) The distribution of signal in tumor and normal tissue (leg muscle (LM), NT, and SG) were compared for the ABY-029 fluorescence and BP. Tumor ROIs were analyzed using: pathologist-determined tumor (entire tumor) and the EGFR-positive pixels within the pathologist-determined tumor. (C) The average signal from each ROI was plotted for all animals. For clarity, only the groups where the means were not statistically different ($p > 0.05$) are shown.

Figure 3. Pixel-by-pixel analysis demonstrates PAI has higher diagnostic accuracy than SAI. (A) ROC curve analysis was performed for each IRDye 700DX, ABY-029, and BP using EGFR IHC
as the gold standard. In all individual mice, BP images have higher AUC than ABY-029 and IRDye 700DX images. (B) Cohort data from each individual tumor group ($n = 6$ or $7$) and All Cell Lines ($n = 20$), also demonstrate that BP maps have higher AUC compared to ABY-029 alone. Cohort statistics generated using the optimum ROC cut off point, and the highest statistic parameter is highlighted in green. BP performance is superior to ABY-029 with the exception of sensitivity and negative predictive value (NPV) in the Det 562 cell line. PPV = positive predictive value, Diag. Acc. = diagnostic accuracy.

**Figure 4.** EGFR heterogeneity is most accurately represented by PAI determined BP. BP and ABY-029 fluorescence images were compared to EGFR IHC on a pixel-to-pixel basis using the Pearson correlation coefficient ($r$). (A) The image pyramid method was used to correct any potential misalignment between pathological and fluorescence images by reducing pixel resolution, which resulted in increased correlation to IHC for all image types. In all four levels of the image pyramid, correlation to EGFR IHC was stronger for BP than ABY-029 or IRDye 700DX. (B) The average and relative change in $r$ were plotted against spatial resolution. Pearson correlation coefficients were determined for Level 2 images (168 μm) of all three tumor lines. A representative scatter plot for each tumor line is shown, in addition to the cohort data at the far right. BP shows a strong positive linear correlation to EGFR IHC in all three tumor groups and is higher than that of ABY-029.

**Figure 5.** Comparison of administration-to-imaging time for PAI and SAI. (A) A representative image of the xenograft tongue tumors are shown for each SAI agent and PAI BP at each time point. The tumor cohort average for each time point is plotted to the right. BP is fairly constant over the
5-hour period, while fluorescence signal decreases for IRDye 700DX and increases for ABY-029 over time. (B) BP map determined CVR is significantly higher ($p < 0.05$) than that of ABY-029 over the first 2 hours after administration, suggesting that BP provides a more stable tumor contrast measure. (C). ROC analysis indicates diagnostic abilities of both imaging methods improve over time; however, BP outperforms ABY-029 images, especially at short administration-to-imaging time points.
Supplementary information

Supplementary Figure 1: Determining doses of paired-agent imaging using auto-fluorescence. A, For each tumor type a representative example of the autofluorescence in a tongue tumor from FaDu, Detroit 562 and A431 cell lines are shown for the 700 and 800 nm channels of the Odyssey CLx scanner and compared to standard H&E and epidermal growth factor receptor (EGFR) immunohistochemistry (IHC). B, The binding potential (BP) calculated for each naïve tongue is shown in the corresponding box plot.

Supplementary Figure 2: A, Fluorescence measurements in tumor and normal tissue are shown in the 700 and 800 channels for auto-fluorescence, 1X dose of ABY-029 and IRDye 700DX (mouse equivalent to the human microdose, 30 nanomoles), and a 1X dose of ABY-029 and 10X dose of IRDye 700DX. B, The naïve and 10X dose values of IRDye 700DX are shown for each tumor type.

Supplementary Figure 3: The effect of the normalization factor (NF) on image contrast. As the normalization ratio increases, the overall signal decreases. This results in a constant contrast-to-variance ratio (CVR) that is stable when NF is varied. However, tumor-to-background ratio (TBR) increases as the signal in the normal tissue (the denominator) decreases. These results indicate that TBR is an unstable measure for binding potential and cannot be used to compare to single agent fluorescence.

Supplementary Figure 4: Image overlays created using imshowpair in Matlab allow visual assessment of spatial alignment between IHC and BP images. In the Checkerboard overlay, IHC is in RGB (brown pixels) while the BP map is a grey scale image.

Supplementary Figure 5: EGFR molecules per cell line for squamous cell carcinomas determined by quantitative flow cytometry. Cell lines used for this study are indicated in green.
Supplementary Table 1: Summary of individual trial and average values of EGFR per cell determined by quantitative flow cytometry.