41. Jacobs, L. Positive margins: The challenge continues for breast surgeons. *Ann. Surg. Oncol.* **15**, 1271–1272 (2008).
42. Jeevan, R. *et al.* Reoperation rates after breast conserving surgery for breast cancer among women in England: Retrospective study of hospital episode statistics. *BMJ* **345**, e4505 (2012).
43. Tichauer, K.M. *et al.* Improved tumor contrast achieved by single time point dual-reporter fluorescence imaging. *J. Biomed. Optics* **17**, 066001 (2012).
44. Tichauer, K.M. *et al. In vivo* quantification of tumor receptor binding potential with dual-reporter molecular imaging. *Mol. Imag. Biol.* **14**, 584–592 (2012).
45. Liu, J.T.C. *et al.* Quantifying cell-surface biomarker expression in thick tissues with ratiometric three-dimensional microscopy. *Biophys. J.* **96**, 2405–2414 (2009).
46. Wang, T.D. *et al.* Functional imaging of colonic mucosa with a fibered confocal microscope for real-time *in vivo* pathology. *Clin. Gastroenterol. Hepatol.* **5**, 1300–1305 (2007).
47. Schmidt, M., Hynes, N.E., Groner, B. & Wels, W. A bivalent single-chain antibody-toxin specific for ERBB-2 and the EGF receptor. *Int. J. Cancer* **65**, 538–546 (1996).
48. Wels, W. *et al.* EGF receptor and p185ERBB-2-specific single-chain antibody toxins differ in their cell-killing activity on tumor cells expressing both receptor proteins. *Int. J. Cancer* **60**, 137–144 (1995).
49. Gaborit, N. *et al.* Time-resolved fluorescence resonance energy transfer (TR-FRET) to analyze the disruption of EGFR/HER2 dimers: A new method to evaluate the efficiency of targeted therapy using monoclonal antibodies. *J. Biol. Chem.* **286**, 11337–11345 (2011).
50. Habib, A.A., Chun, S.J., Neel, B.G. & Vartanian, T. Increased expression of epidermal growth factor receptor induces sequestration of extracellular signal-related kinases and selective attenuation of specific epidermal growth factor-mediated signal transduction pathways. *Mol. Cancer Res.* **1**, 219–233 (2003).
51. Moasser, M.M., Basso, A., Averbuch, S.D. & Rosen, N. The tyrosine kinase inhibitor ZD1839 ("iressa") inhibits HER2-driven signaling and suppresses the growth of HER2-overexpressing tumor cells. *Cancer Res.* **61**, 7184–7188 (2001).
52. Van De Sompel, D., Garai, E., Zavaleta, C. & Gambhir, S.S. A hybrid least squares and principal component analysis algorithm for Raman spectroscopy. *PLoS ONE* **7**, e38850 (2012).
53. Chithrani, B.D., Ghazani, A.A. & Chan, W.C.W. Determining the size and shape dependence of gold nanoparticle uptake into mammalian cells. *Nano Lett.* **6**, 662–668 (2006).

SUPPLEMENTARY INFORMATION

Demultiplexing SERS spectra using direct classical least squares (DCLS)

This study employs a demultiplexing method described previously[1]. In a realistic spectral measurement, pure spectra from single- or multi-flavor particle mixtures are mixed with varying magnitudes of broadband background signal (mainly due to laser background and autofluorescence background) and zero-mean Gaussian-distributed white noise (includes shot noise and other stochastic noise sources such as detector readout noise and dark counts). Other than sources of noise, it is assumed that each measured spectrum consists of a weighted sum of fixed nanoparticle spectra (F_n) and broadband background signals (B). Based on the assumption that the combination is linear, we employ a linear least-squares algorithm to compute the relative nanoparticle weights (w_n). A third-order polynomial is included to account for broadband background signals that are not captured by the reference spectrum (B).

$$S = \sum_n w_n F_n + kB + \sum_m a_m P_m + R \qquad (1)$$

where:

$\boldsymbol{S}$ = measured spectral data
w_n = weight of SERS flavor n
F_n = known reference spectrum of SERS nanoparticle flavor n
k = scaling factor for background signal magnitude
B = known reference spectrum of broadband background
a_m = weight of mth-order polynomial term
$\boldsymbol{P}_m$ = mth-order polynomial term (for baseline correction)
$\boldsymbol{R}$ = residual (minimized by least-squares algorithm)

In order to process raw spectral data, we first subtract the constant detector background that is obtained with the laser illumination turned off. This background is due to detector noise, which is primarily electronic read noise and a smaller amount of dark thermal noise (dependent upon integration time and the level of detector cooling). The spectra are cropped and only the range 700–2000 cm^{-1} (Raman shift) is used for demultiplexing analysis.

Reference spectra are obtained for pure NP flavors, F_n, as well as for the broadband background, B. In order to obtain a NP reference spectrum, a 3–5 μL drop of NPs (800 pM concentration) is placed on the surface of the spectral-detection device, or at a fixed working distance away from the device. A spectrum is recorded from the NP sample. Next, another spectrum is recorded from a drop of the buffer in which the NPs were suspended (generally water), under the exact same conditions (*e.g.* droplet volume, detection geometry, detector settings). These two spectra are subtracted from each other to obtain the pure reference spectrum of the NPs. In order to obtain a background reference spectrum, B, the sample (cells or tissue) is stained with an appropriate buffer solution (generally PBS) that is devoid of NPs. A spectrum of the buffer-stained sample is acquired under the exact same conditions (*e.g.* laser and detector settings, measurement geometry, etc.) as is used for measuring NP-stained samples. In certain cases, multiple spectra are acquired to obtain an average background reference spectrum, B. Alternatively, a principle component analysis (PCA) may be performed to determine the most appropriate reference spectrum or spectra (if more than one background reference spectrum is desired).

Ratiometric quantification of specific vs. nonspecific binding

For ratiometric quantification of specific vs. nonspecific binding, it is assumed that negative-control NPs exhibit identical nonspecific behavior to targeted NPs. In this study, targeted NP flavors are conjugated to monoclonal antibodies (anti-EGFR or anti-HER2) and a negative-control NP flavor is conjugated to an appropriate isotype-control antibody. NP characteristics such as charge, size, and overall chemistry are therefore very similar between NPs, except for the antigen-binding sites of the targeted mAbs. With these assumptions, the measured weight, w_n, of each NP flavor, as determined by least-squares demultiplexing, can be modeled as:

$$w_n = g_n \cdot I_n \cdot c_n \cdot \sigma_n \qquad (2)$$

where:

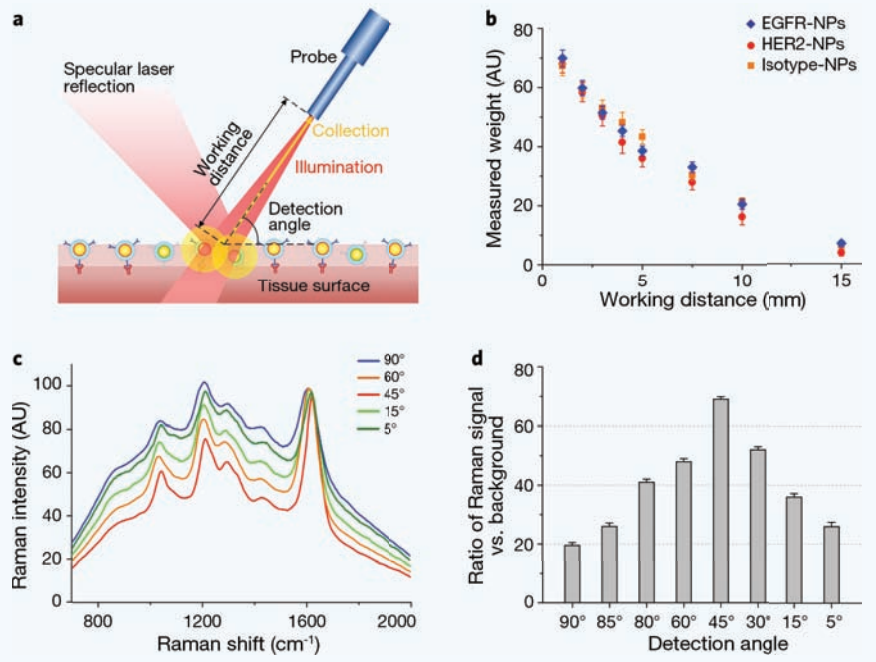
g_n = electronic and system gain (detector quantum yield, optical throughput, etc.)
I_n = laser power
c_n = concentration of NP flavor n
σ_n = average scattering cross section of NP flavor n

The ratiometric strategy seeks to determine the true concentration ratio between NP flavors:

$$\frac{c_1}{c_2} = \frac{w_1}{w_2} \cdot \frac{g_2 I_2 \sigma_2}{g_1 I_1 \sigma_1} = \frac{w_1}{w_2} \cdot \kappa \qquad (3)$$

The least squares demultiplexing algorithm yields the weight ratio, w_1/w_2. We assume that the ratio $g_2I_2\sigma_2/g_1I_1\sigma_1$ is a constant (κ) and can be determined *via* a calibration measurement of a known concentration ratio of NPs, typically an equimolar ratio where $c_1/c_2 = 1$.
This assumption is based on the following arguments:

1. Under ideal demultiplexing conditions, in which there is negligible crosstalk between NP flavors in the least-squares algorithm, the measured weights, w_n, are linear and directly proportional to the actual NP concentrations, c_n.
2. A single laser illumination source is used to illuminate all NPs identically within each tissue region, such that $I_n = I_1 = I_2$ for all NP flavors.
3. The electronic and system gain are either constant and identical for each NP flavor (*e.g.* optical throughput) or the ratio of these



Supplementary Figure 1 Optimization experiments of a miniature hand-held contact probe for Raman detection. (**a**) Schematic illustration of the detection of NPs on tissue surfaces. (**b**) Effect of working distance on demultiplexed NP weights (signal intensity) with the handheld probe (n = 5). The probe has a detection angle of 45° with a variable working distance ranging from 1 mm to 15 mm. (**c**) Raw spectra as a function of the detection angle. All spectra are normalized by the magnitude of the main peak at ~1620 cm^{-1}. (**d**) Ratio of Raman signal (from NPs) versus background signal (from tissue autofluorescence and laser reflections).

constants (g_2/g_1) are constant due to a linear behavior of these constants with respect to perturbations (*e.g.* fluctuations in detector quantum yield due to temperature changes, variations in signal due to changes in detector integration time, and/or variations in signal due to changes in measurement geometry). In practice, all calibration measurements are performed with the exact same detector and system settings as the actual experiments, and the individual gains g_n can be treated as constants. However, this is not strictly required, since a single detector and optical detection path are used to record all NP signals and the ratio g_2/g_1 remains constant even under variable gain settings.

4. The ratio of the scattering cross sections σ_2/σ_1 is a constant due to the fact that the Raman-active layer in the SERS NPs is encapsulated in silica and shielded from environmental factors. Any perturbation in the scattering cross sections — due to optical, thermal, or other effects — is assumed to affect all NP flavors in a linear fashion such that σ_2/σ_1 remains constant. In practice, all calibration measurements are performed under the same optical and environmental conditions as the actual experiments and the individual cross sections σ_n can be treated as constants. However, this is not strictly required.

Note that with targeted SERS NPs, it is possible that as the NPs diffuse through tissues, the differential binding between various NP flavors may alter the initial ratio of the NP mixtures within the staining "reservoir" of NPs. Therefore, the initial 1:1 ratio between all NP flavors in the staining solution may not be true deeper within tissues. However, due to the large size of our SERS NPs (~120 nm in diameter), the NPs are unable to diffuse well within tissues, and are largely confined to the superficial surface when topically applied on tissues[2].

(*Ex vivo* and *in vivo*) optimization of a miniature hand-held probe for contact detection of SERS NPs on tissues

One challenge in using a hand-held probe for NP detection is that the working distance and detection angle (see **Supplementary Fig. 1a**) are difficult to control. In particular, for *in vivo* experiments, uneven tissue surfaces and motion from respiration can cause fluctuations in working distance and detection angle. While our ratiometric quantification strategy is designed to mitigate these effects, large variations in signal strength (SNR) and signal-to-background ratio (SBR) can reduce the accuracy of quantitive measurements of NPs located on tissue surfaces. Therefore, here, we have used contact detection to minimize working-distance and angle-dependent effects. The probe consists of a prism to fix the working distance and detection angle, where experiments were performed to determine the optimal working distance and detection angle. For these measurements, mice were anesthetized *via* intraperitoneal injection of ketamine and xylazine, and a small flat area of tissue was stained with an equimolar mixture of 3 flavors of conjugated NPs (200 pM per flavor).

Supplementary Fig. 1b plots the variation of demultiplexed NP weights as a function of working distance (1–15 mm at a 45° detection angle), showing that signal strength degrades with working distance and that a contact probe (zero working distance) yields the highest detection sensitivity. These results are expected due to the reduction in the numerical aperture (NA) of collection with working distance. **Supplementary Fig. 1c** shows raw spectra collected under different detection angles and **Supplementary Fig. 1d** plots the ratio of Raman signal versus tissue background. The data suggest that detection at 45° provides an optimal SBR, likely due to the reduced collection of specularly reflected laser light. Based on these results, a 2.5-mm 45° prism was adhered to the probe tip to fix the detection angle at ~45° for *ex vivo* and *in vivo* experiments.

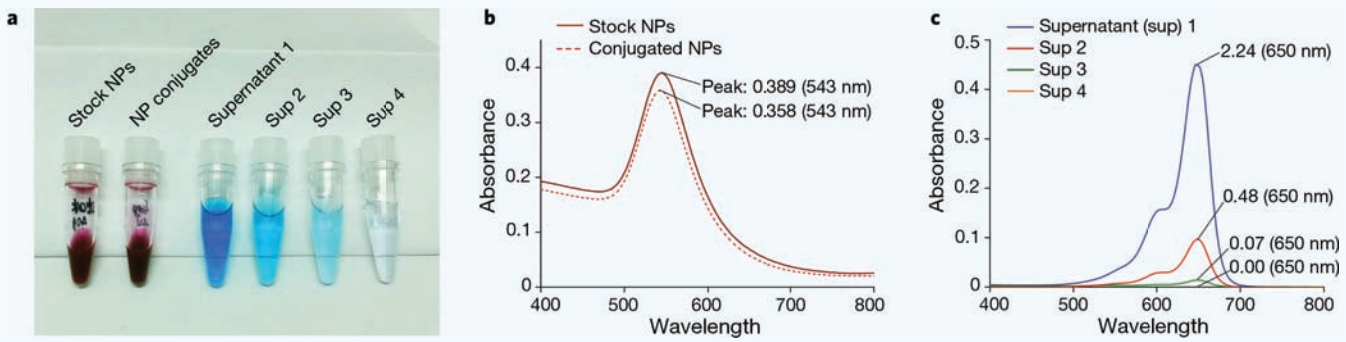
Optimization of conjugation and staining parameters

Supplementary Table 1 Flow-cytometry fluorescence ratio of HER2-NPs vs. isotype-NPs on SkBr3 cells under different conjugation and staining conditions.

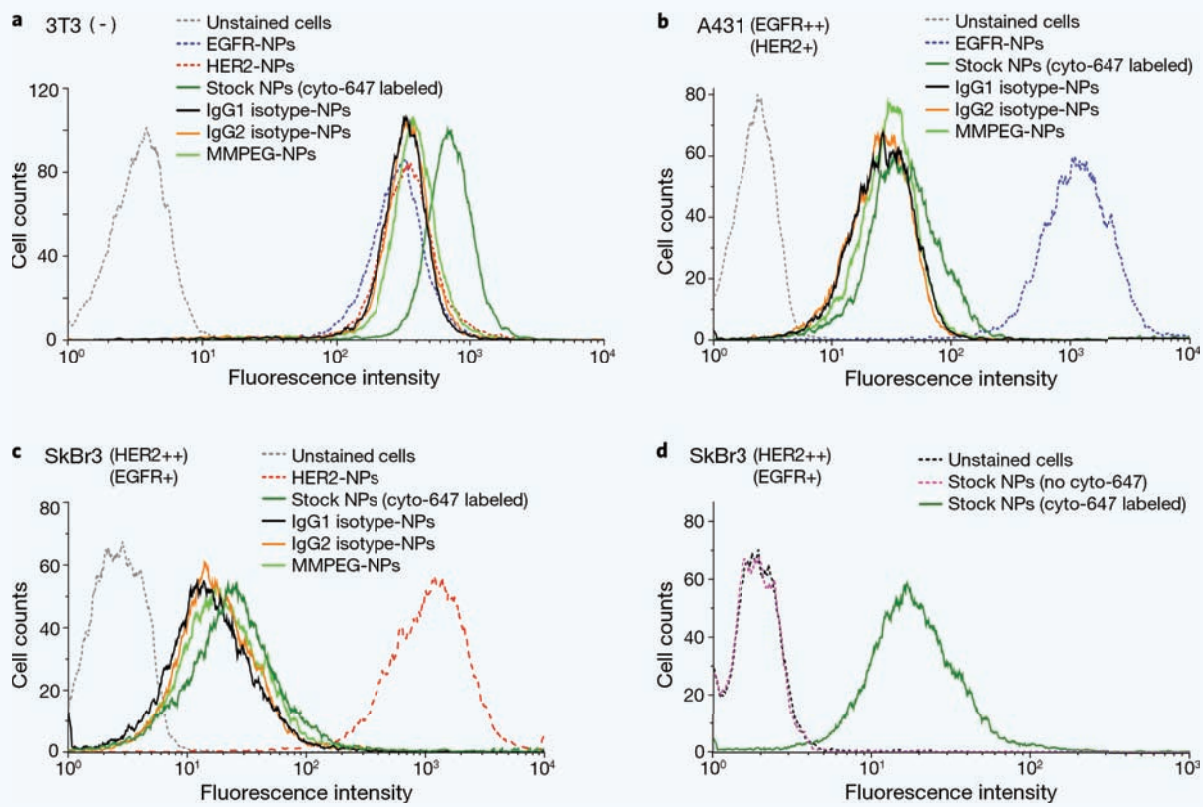
Antibody/NP ratio	$MM(PEG)_{12}$ amount	Staining conc. on cells (pM of NPs)	Staining time on cells (min)	FI ratio
350	0	40	15	9
350	6×10^6/NP	40	15	11
350	6×10^6/NP	100	15	21
350	6×10^6/NP	100	60	14
350	6×10^6/NP	200	15	49
420	6×10^6/NP	200	15	63
500*	6×10^6/NP	200	15	92

*For antibody/NP ratios greater than 500, aggregation of the conjugated NPs becomes increasingly apparent.

Calculating the concentration of NP conjugates *via* UV-VIS spectrophotometry

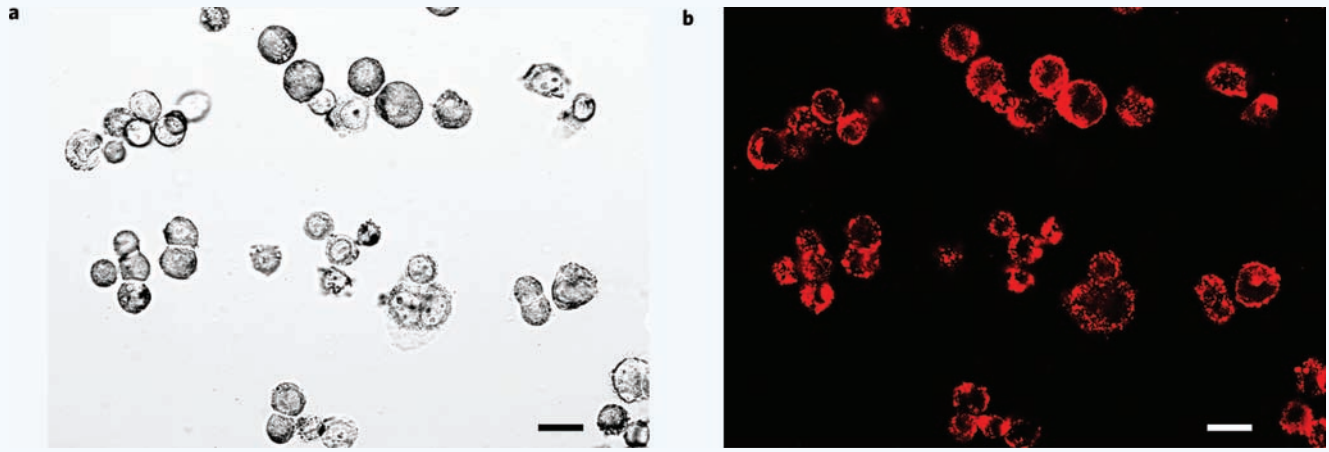


Supplementary Figure 2 UV-VIS spectrophotometry for purifying and calculating the concentration of NP conjugates. (**a**) The color of NPs and supernatants during purification; (**b**) absorbance spectra of stock NPs and conjugated NPs for calculation of NP concentration (50× dilution); (**c**) absorbance spectra of supernatant (3× dilution; the main component is cyto 647) from 4 rounds of purification.

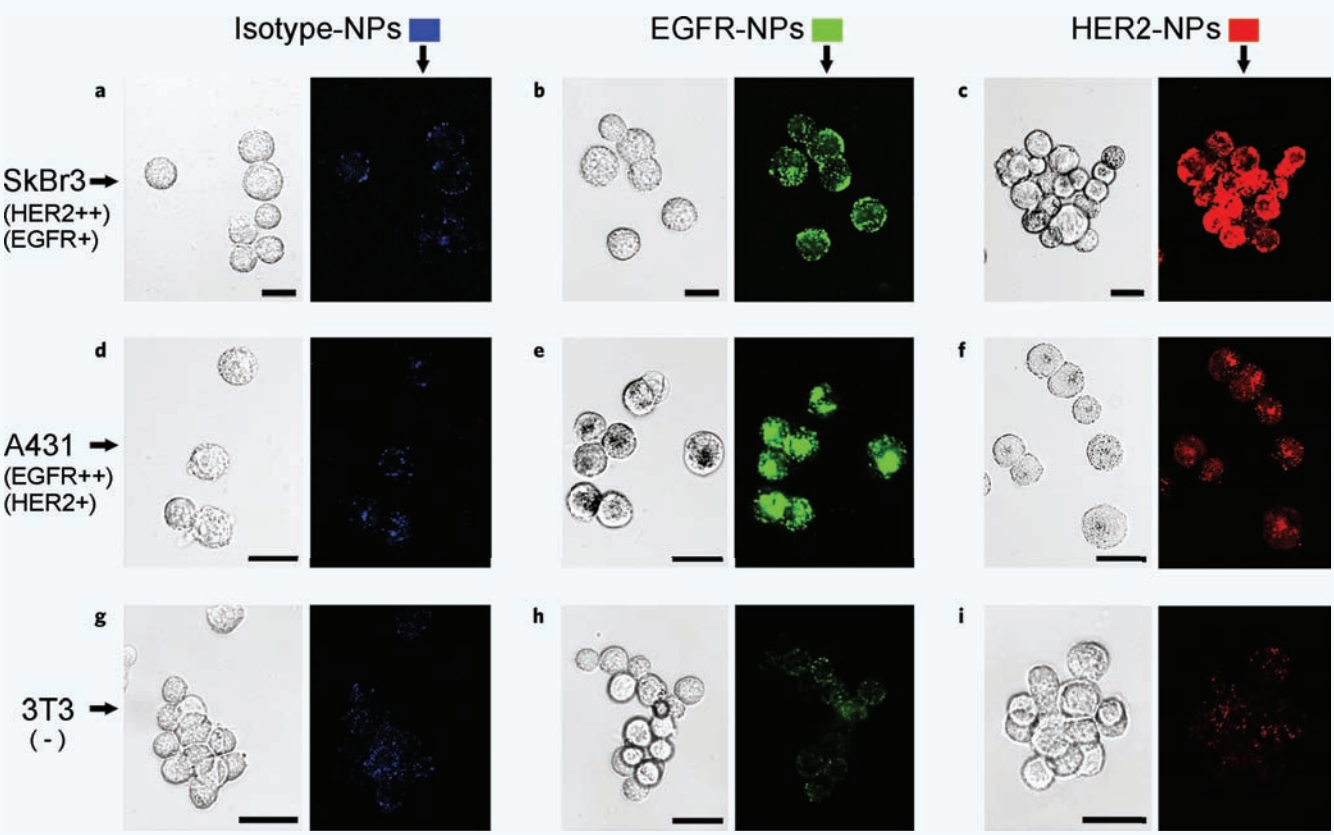


Supplementary Figure 3 Flow cytometry comparison of control NPs with cultured cells. For (**a–c**), four types of control NPs, labelled with the fluorophore cyto-647, were prepared for this optimization experiment, including: (1) stock NPs that were only labelled with cyto-647, (2) MMPEG-NPs that were conjugated with an excess of $MM(PEG)_{12}$ to block thiols on the NP surface, and (3) IgG1 and (4) IgG2 isotype-NPs that were conjugated to mouse IgG1 isotype mAb and human IgG2 isotype mAb, respectively, along with $MM(PEG)_{12}$ (refer to the conjugation protocol). NPs conjugated with anti-EGFR and anti-HER2 monoclonal antibodies were also prepared with the equivalent fluorophore: NP ratio and antibody: NP ratio as the isotype-NPs. The various NPs were individually used to stain (**a**) 3T3 (–), (**b**) A431 (EGFR++, HER2+) and (**c**) SkBr3 (HER2++, EGFR+) cell lines for 15 min (200 pM), followed by flow cytometry analysis. The flow experiment shown in panel a used a higher laser power to compare the behaviors of different control NPs. Although previous reports describe the use of pure (stock) SERS NPs or PEG-modified NPs[3,4] as controls, the data in panel a suggest that IgG1 and IgG2 isotype-NPs more accurately mimic the nonspecific behavior of EGFR-NPs and HER2-NPs on the control cell line (3T3). In addition, panels b and c show the fluorescence ratio of targeted NPs vs. various control NPs for the tumor cell lines A431 and SkBr3. (**d**) Fluorescence from unstained cells and NP-stained cells. Non-fluorescent stock NPs (S421) and cyto-647 labeled stock NPs (S421) were individually used to stain SkBr3 cell lines for 15 min, followed by flow cytometry analysis. The results show that since the NPs are encapsulated in silica, there are no surface-enhanced fluorescence effects and that the fluorescence signals are only due to the cyto-647 fluorophores that are conjugated to the NPs.

Observation of NP distribution on cells *via* confocal microscopy

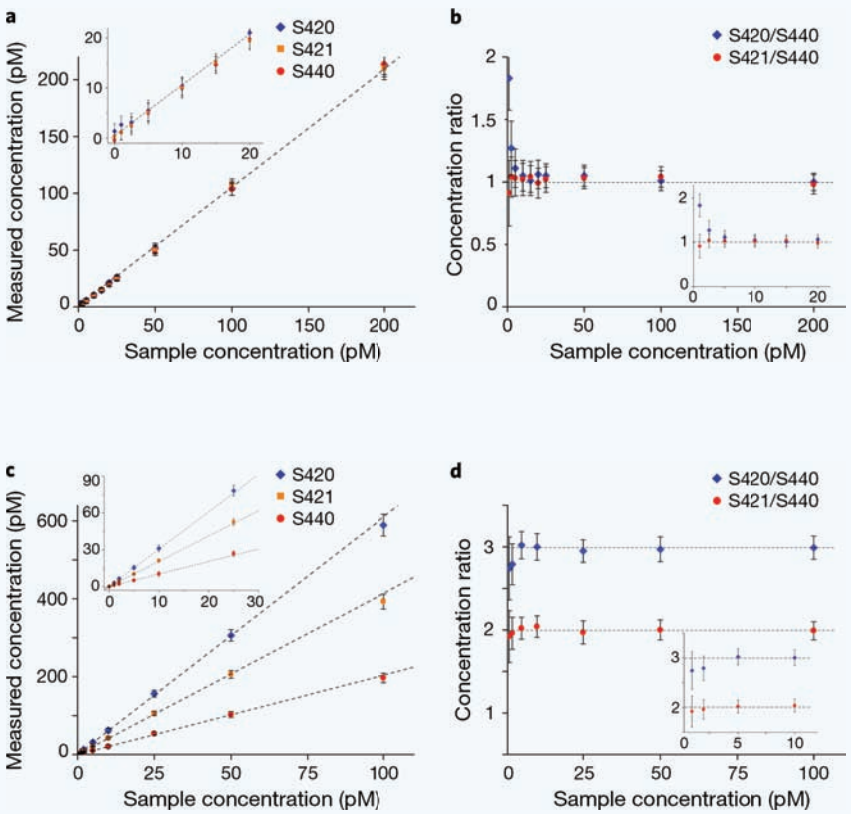


Supplementary Figure 4 Confocal microscopy of the NP distribution on cells using a reduced pinhole, which provides a thin optical section through the center of the cells. (**a**) Bright-field image; (**b**) fluorescent image. The scale bars are 20 μm. The fluorescence signal is primarily localized to the periphery of the cells, indicating that the NPs are mostly bound to the cell surface.



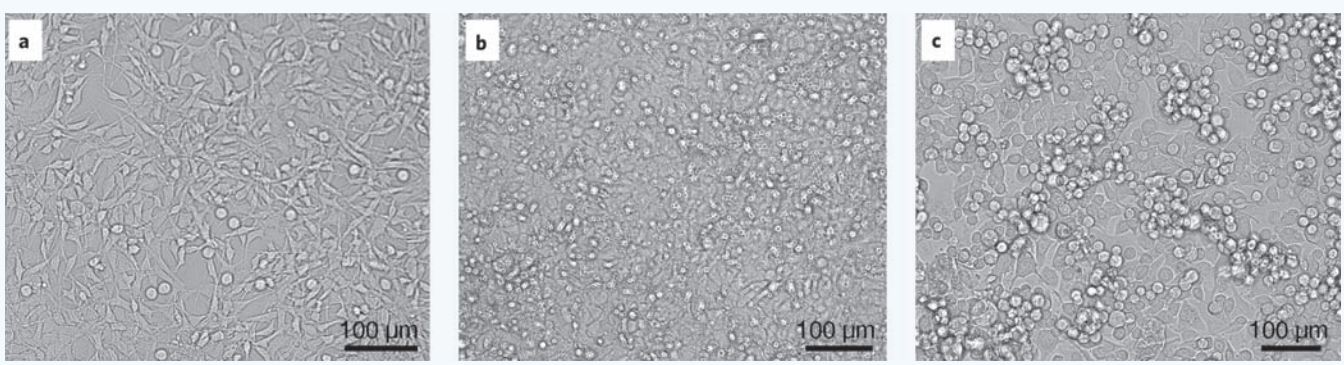
Supplementary Figure 5 Confocal microscopic images of different cells stained with conjugated NPs. These cells are in suspension rather than adhered to a culture dish as in **Supplementary Fig. 7**. The left side of each panel is a bright-field image, and the right side is a fluorescence image (illumination 633 nm; collection 660–700 nm). The scale bars all represent 20 μm. For these images, the confocal pinhole was set to a large diameter to image entire cells, including all NPs attached to each cell. Weak fluorescence is seen from SkBr3 and A431 cells labeled with isotype-NPs as well as from all 3T3 (negative-control) samples (**a,d,g–i**). SkBr3 and A431 cells labeled with EGFR-NPs and HER2-NPs show strong fluorescence intensities (**b,c,e,f**) and provide visual verification of the flow cytometry results.

(*In vitro*) limit-of-detection test

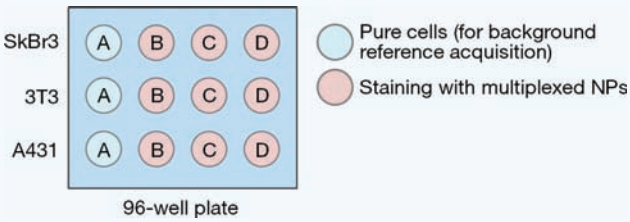


Supplementary Figure 6 Limit-of-detection (linearity) test for a 3-flavor mixture. (**a,b**). Three NP flavors were mixed in an equimolar ratio (1:1:1) and diluted to different concentrations (1, 2.5, 5, 10, 15, 20, 25, 50, 100, and 200 pM). (**c,d**). Three NP flavors were mixed in a ratio of 3:2:1 (S420:S421:S440) and diluted to different concentrations (S440: 1, 2, 5, 10, 25, 50, 100, and 200 pM). The horizontal axis in panels (**c**) and (**d**) indicates the concentration of S440 in the NP mixtures. For each mixture, a 50-μL volume was transferred to one well of a 96-well plate, and 3 positions within each well were scanned twice with a detection probe, for a total of 6 spectra (3 positions and 2 spectral acquisitions per position). The spectra were demultiplexed to calculate the weights of each flavor of NP. In order to calculate the concentration of NPs based on the measured weights, a calibration measurement was performed by recording the weights of S420, S421 and S440 measured from a stock mixture. The concentration ratios of S420/S440 and S421/S440 were calculated and plotted in (**b**) and (**d**). The measured concentrations and concentration ratios show good linearity from 2.5 to 200 pM when 3 flavors of NPs are multiplexed, with larger errors in terms of concentration (>20%) and concentration ratio (>15%) appearing below 2.5 pM. Therefore, all data presented in our manuscript correspond to NP concentrations larger than 2.5 pM (>4 pM for cell monolayer experiments and >10 pM for tissue experiments).

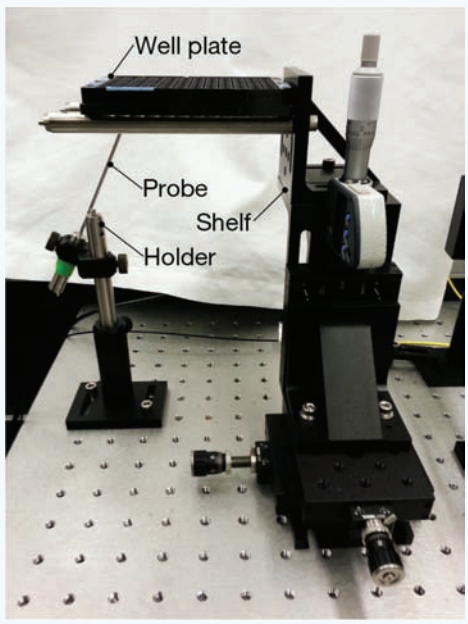
(*In vitro*) cell monolayer experiments with 96-well plate



Supplementary Figure 7 Microscopic images of confluent cell monolayers of (**a**) 3T3, (**b**) A431 and (**c**) SkBr3 in one well of a 96-well plate.

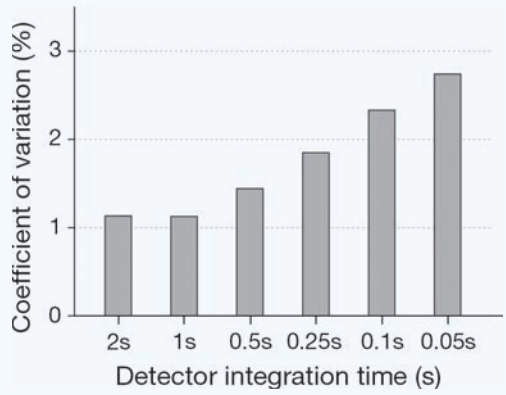


Supplementary Figure 8 Schematic of cell-monolayer experiments on a 96-well plate. For each cell line, cells were counted and seeded into four wells A–D (1×10^5 cells in 50 µL of media per well for A431 and SkBr3, and 0.5×10^5 cells in 50 µL of media per well for 3T3), followed by incubation for 24 h to obtain similar cell monolayers (about 2×10^5 cells). Prior to experiments, media was removed and the cells were washed with PBS. In well A, PBS was added to allow for the acquisition of a background measurement in the absence of NPs. For the 3 remaining wells of each cell line (B–D), an equimolar mixture of targeted (anti-EGFR and/or anti-HER2) and nontargeted (isotype control) NPs was added to allow for ratiometric analyses of specific vs. nonspecific binding of the NPs. Both 2-flavor and 3-flavor mixtures were used at a staining concentration of 150 pM per flavor and a total staining volume of 100 µL. Staining was performed for 15 min followed by up to three rounds of rinsing with PBS. Spectral measurements and ratiometric analysis was performed before and after each rinse step.



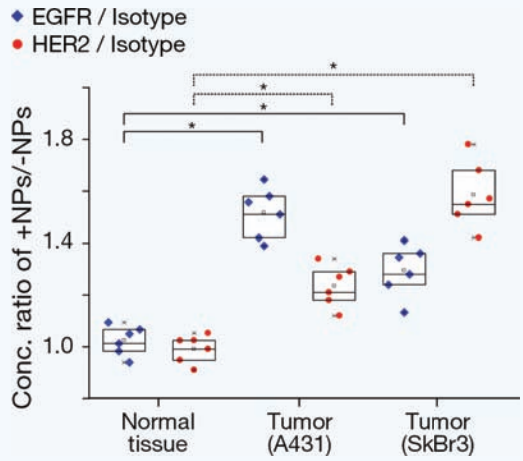
Supplementary Figure 9 Experimental platform for the spectral detection of cell monolayers on a well plate. A probe holder and a shelf were used to maintain the working distance and the illumination and detection angles (probe angle) during measurements.

(*In vivo*) Effect of detector integration time on measurement uncertainty



Supplementary Figure 10 Coefficient of variation of the measured concentration of NPs on tissues under different integration times. For each data set, 50 acquisitions were taken.

(*In vivo*) Effect of NP flavor on ratiometric analysis



Supplementary Figure 11 The concentration ratio of targeted and nontargeted NPs topically applied on exposed tissues and measured *in vivo*. Tumors and normal tissues were topical stained with a mixture of 3 NP flavors (three measurements on each of two mice for a total n = 6). In these experiments, the NP flavors utilized for each targeted agent are different from those in **Fig. 6e**. For the experiments in **Fig. 6e**, S420 NPs were conjugated to an anti-HER2 mAb and S421 NPs were conjugated to an isotype-control mAb. Here, S421 NPs are conjugated to an anti-HER2 mAb and S420 NPs are conjugated to an isotype-control mAb. *P-value < 0.001.

SUPPLEMENTARY REFERENCES

1. Leigh, S.Y., Som, M. & Liu, J.T.C. Method for assessing the reliability of molecular diagnostics based on multiplexed sers-coded nanoparticles. *PLoS ONE* **8**, e62084 (2013).
2. Zavaleta, C.L. *et al.* Preclinical evaluation of Raman nanoparticle biodistribution for their potential use in clinical endoscopy imaging. *Small* **7**, 2232–2240 (2011).
3. Jokerst, J.V., Miao, Z., Zavaleta, C., Cheng, Z. & Gambhir, S.S. Affibody-functionalized gold–silica nanoparticles for Raman molecular imaging of the epidermal growth factor receptor. *Small* **7**, 625–633 (2011).
4. Wang, Y.Q., Yan, B. & Chen, L.X. SERS tags: Novel optical nanoprobes for bioanalysis. *Chem. Rev.* **113**, 1391–1428 (2013).