University of Groningen

Zr-89-atezolizumab imaging as a non-invasive approach to assess clinical response to PD-L1 blockade in cancer

Bensch, Frederike; van der Veen, Elly L.; Lub-de Hooge, Marjolijn N.; Jorritsma-Smit, Annelies; Boellaard, Ronald; Kok, Iris C.; Oosting, Sjoukje F.; Schröder, Carolina P.; Hiltermann, T. Jeroen N.; van der Wekken, Anthonie J.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2018

Link to publication in University of Groningen/UMCG research database

89Zr-atezolizumab imaging as a non-invasive approach to assess clinical response to PD-L1 blockade in cancer

Frederike Bensch[1], Elly L. van der Veen[1], Marjolijn N. Lub-de Hooge[2,3], Annelies Jorritsma-Smit[2], Ronald Boellaard[3], Iris C. Kok[1], Sjoukje F. Oosting[1], Carolina P. Schröder[1], T. Jeroen N. Hiltermann[4], Anthonie J. van der Wekken[4], Harry J. M. Groen[4], Thomas C. Kwee[3], Sjoerd G. Elias[5], Jourik A. Gietema[1], Sandra Sanabria Bohorquez[6], Alex de Crespigny[6], Simon-Peter Williams[6], Christoph Mancao[7], Adrienne H. Brouwers[3], Bernard M. Fine [ID][6] and Elisabeth G. E. de Vries [ID][1]*

Programmed cell death protein-1/ligand-1 (PD-1/PD-L1) blockade is effective in a subset of patients with several tumor types, but predicting patient benefit using approved diagnostics is inexact, as some patients with PD-L1-negative tumors also show clinical benefit[1,2]. Moreover, all biopsy-based tests are subject to the errors and limitations of invasive tissue collection[3–11]. Preclinical studies of positron-emission tomography (PET) imaging with antibodies to PD-L1 suggested that this imaging method might be an approach to selecting patients[12,13]. Such a technique, however, requires substantial clinical development and validation. Here we present the initial results from a first-in-human study to assess the feasibility of imaging with zirconium-89-labeled atezolizumab (anti-PD-L1), including biodistribution, and secondly test its potential to predict response to PD-L1 blockade (ClinicalTrials.gov identifiers NCT02453984 and NCT02478099). We imaged 22 patients across three tumor types before the start of atezolizumab therapy. The PET signal, a function of tracer exposure and target expression, was high in lymphoid tissues and at sites of inflammation. In tumors, uptake was generally high but heterogeneous, varying within and among lesions, patients, and tumor types. Intriguingly, clinical responses in our patients were better correlated with pretreatment PET signal than with immunohistochemistry- or RNA-sequencing-based predictive biomarkers, encouraging further development of molecular PET imaging for assessment of PD-L1 status and clinical response prediction.

Excitement about durable responses in cancer patients has spurred clinical investigations and marketing approvals for immunotherapies based on checkpoint blockade of PD-1 and its ligand PD-L1. Identifying patients likely to benefit from these therapies, however, remains challenging. Two diagnostic tests based on PD-L1 immunohistochemistry (IHC) have been approved to predict patient benefit. However, not all patients with high tumor PD-L1 benefit from treatment with checkpoint inhibitors, and some with no PD-L1 staining also show benefit[1,2]. At present, it is unclear

whether this is primarily due to artifacts related to limited tissue sampling or under-appreciated facets of PD-L1 biology, including spatial and temporal heterogeneity[14–20]. Other predictive biopsy-based biomarkers have been evaluated but are also subject to errors and limitations of invasive tissue collection[3–11]. As suggested by preclinical reports, a macroscopic, non-invasive molecular imaging readout for PD-L1 could provide new insights by assessing the PD-L1 status throughout the whole body, potentially at multiple time points, thus capturing information about the tumor immune infiltrate and its response to therapy[12,13]. Such insights may be important in optimizing the use of existing treatments and in developing new immunotherapeutic agents and combinations.

Here we present results from a first-in-human imaging study with 89Zr-labeled atezolizumab (anti-PD-L1). We enrolled 25 patients with locally advanced or metastatic bladder cancer, non-small cell lung cancer (NSCLC), or triple-negative breast cancer (TNBC) between March and November 2016 (Supplementary Table 1). Three patients discontinued prematurely, owing to disease progression, before tracer injection or during imaging procedures. Twenty-two patients completed the full imaging series of up to four PET scans and were subsequently treated with atezolizumab until progressive disease (PD).

89Zr-atezolizumab injection was safe, with only one related low-grade adverse event (Supplementary Table 2). The side effects of atezolizumab monotherapy were comparable to those in previous reports, except for the higher incidence of mainly low-grade infusion-related reactions ($n = 4$, including one grade 3 event)[4,8,21].

We added 10 mg unlabeled atezolizumab to the tracer to prevent rapid clearance during imaging. Pharmacokinetic analysis showed good correlation with activity of the blood pool on PET and confirmed that the circulating tracer dose corresponded with a serum atezolizumab concentration reached with 0.1–0.3 mg kg⁻¹ atezolizumab, which is almost 100-fold lower than that reached with the recommended atezolizumab treatment dose (Supplementary Fig. 1a,b)[22]. Because the day 4 blood pool, liver, and kidney 89Zr-atezolizumab uptake was comparable to the results of other 89Zr

[1]Department of Medical Oncology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands. [2]Clinical Pharmacy and Pharmacology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands. [3]Medical Imaging Center, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands. [4]Pulmonary Oncology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands. [5]Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands. [6]Genentech, San Francisco, CA, USA. [7]Genentech, Basel, Switzerland. *e-mail: e.g.e.de.vries@umcg.nl

antibody tracers with well-known kinetics over time[23–26], and tumor lesions were visualized satisfactorily, we considered this unlabeled antibody dose to be sufficient. Moreover, because PET scans on days 0 and 2 did not add valuable information in the first ten patients, we decided to continue with scans on days 4 and 7 only.

First, we assessed [89]Zr-atezolizumab biodistribution. We observed low uptake in the healthy brain, subcutaneous tissue, muscle, compact bone, and lung, and higher uptake over time in the intestines, kidney, and liver; these last three probably reflect antibody metabolism and elimination (Fig. 1a and Supplementary Fig. 2). [89]Zr-atezolizumab uptake also increased slightly in bone marrow over time. This uptake was considered specific and not due to accumulation of free [89]Zr, as uptake in compact bone was low and stable over time, and the tracer remained intact in serum (Fig. 1a and Supplementary Fig. 3a,b). Non-malignant lymph nodes were also visualized with [89]Zr-atezolizumab PET in the majority of patients on days 4 and 7 (Fig. 1b). In contrast to reports about [89]Zr antibody tumor cell and growth factor targeting tracers[23–26], we observed increasing high and variable [89]Zr-atezolizumab uptake in the spleen, compatible with target-specific binding (Fig. 1a). PD-L1 IHC showed variable expression in non-malignant lymph nodes and prominent expression in the spleen, the latter coinciding with CD8 expression (Fig. 1c,d). On the basis of morphology, this is mainly attributed to endothelial littoral cells, which line the venous sinusoids (Fig. 1d). These cells are CD8α- and PD-L1-positive and stain for CD68, a protein highly expressed by macrophages, suggesting a relationship between these two cell types (data not shown)[27]. Overall, the observed [89]Zr-atezolizumab uptake in lymphoid tissue might serve as a surrogate for the activation state of the body's immune system or as a measure for abundant PD-L1 expression. Finally, sites of clinically observed inflammation in individual patients were identified on [89]Zr-atezolizumab PET (Supplementary Fig. 4). At these sites, depending on the phase of inflammation, different PD-L1-expressing immune cells can be found[28], which might be visualized with PET.

Furthermore, we were able to visualize lesions at all main metastatic sites (Fig. 2a,b). Because patients with central nervous system metastases were excluded from the study, it is unknown whether these metastases could also be visualized. The maximum standardized uptake value (SUVmax) of tumor lesions (overall and according to tumor type), tumor-to-background ratio for lung and bone metastases, and tumor-to-blood ratio increased over time, with the first
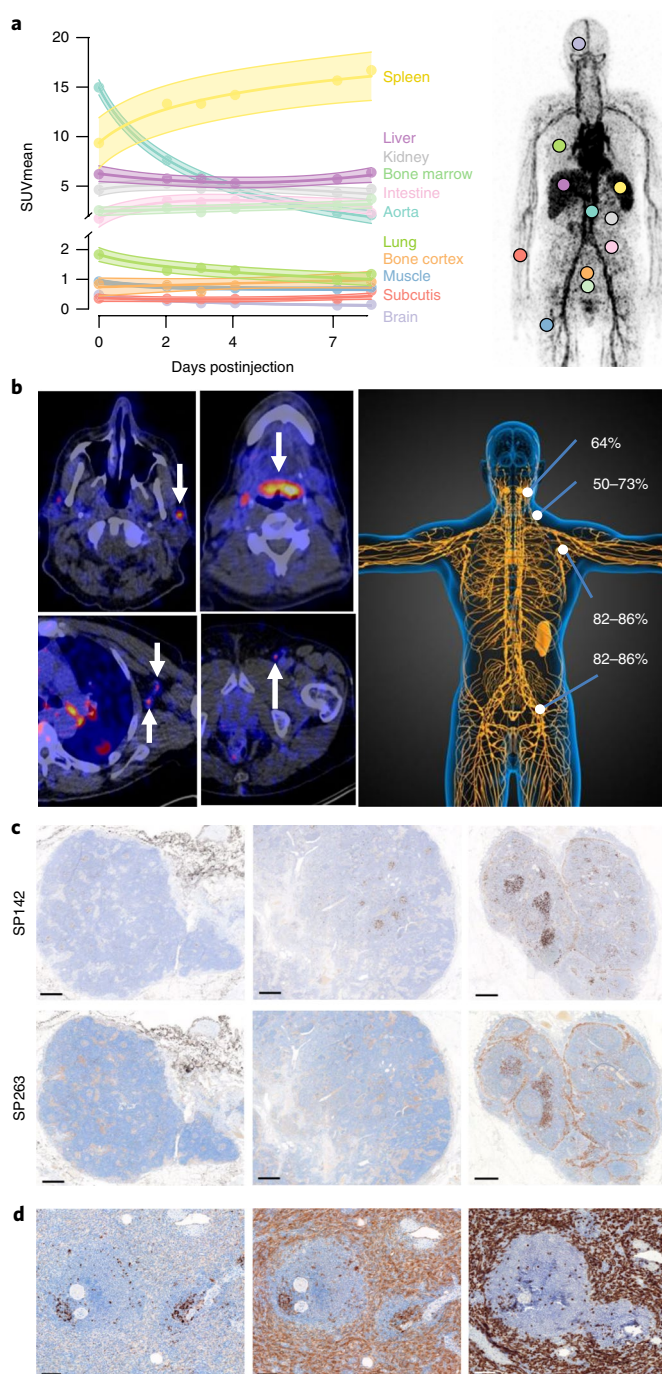

Fig. 1 | [89]Zr-atezolizumab biodistribution and PD-L1 IHC in healthy tissue. **a**, Tracer uptake as mean SUVmean (95% CI) per time point in healthy tissue and blood 1h after tracer injection (day 0) and on days 2, 4, and 7 postinjection (±1 day); measured in ten patients on days 0 and 2 (±1 d) and in all 22 patients on days 4 and 7 (±1 d); fitted regression lines with 95% CI based on linear mixed effect models (764 data points from 12 localizations (liver uptake was measured in two regions per patient per time point) in 22 patients; 4 missing data points). Representative maximum intensity projection of a [89]Zr-atezolizumab PET scan on the right indicates location of measured healthy tissue. **b**, Percentage of patients (n = 22) with [89]Zr-atezolizumab uptake in healthy lymphoid tissue on day 7 (right) and examples of [89]Zr-atezolizumab uptake in the Waldeyer's tonsillar ring (upper middle) and small normal lymph nodes in the neck (upper left), axillary region (lower left), and inguinal region (lower middle) on day 7 postinjection (PET scans were performed once per patient and time point). **c**, Examples of three normal non-malignant lymph nodes with PD-L1 IHC (SP142, upper panel; SP263, lower panel) illustrating heterogeneous PD-L1 staining within and between samples (spare tissue, not obtained from study population) (scale bars, 500 μm; IHC was performed once). **d**, Healthy spleen with a lymphoid follicle surrounded by endothelial littoral cells with intense PD-L1 (SP142, left; SP263, middle) and high CD8 (right) staining (scale bars, 100 μm; IHC was performed once).

two stabilizing at day 7 postinjection (Supplementary Fig. 5a–c). We further report day 7 uptake data, as the SUVmax values of days 4 and 7 were highly correlated (Supplementary Fig. 5d), and the tumor-to-blood ratio was the most favorable on the latest scan moment. Tumor [89]Zr-atezolizumab uptake was generally high (Fig. 2b), with an overall geometric mean SUVmax of 10.4 (95% confidence interval (CI) 8.5–12.7; range 1.6–46.1). We observed major within-patient SUVmax heterogeneity in the 20 patients with more than one lesion, with a median fold difference of 2.2 (range 1.0–9.4) and a median coefficient of variation of 12.2% (range 0.7–39.3%). Tracer uptake varied per site of metastases ($P = 2.2 \times 10^{-7}$): in sites with at least ten observations, liver metastases had the highest uptake and lung metastases the lowest (Fig. 2c). Moreover, tumor tracer uptake differed per tumor type ($P = 0.016$), with TNBC showing on average 50% (95% CI 17–70%)
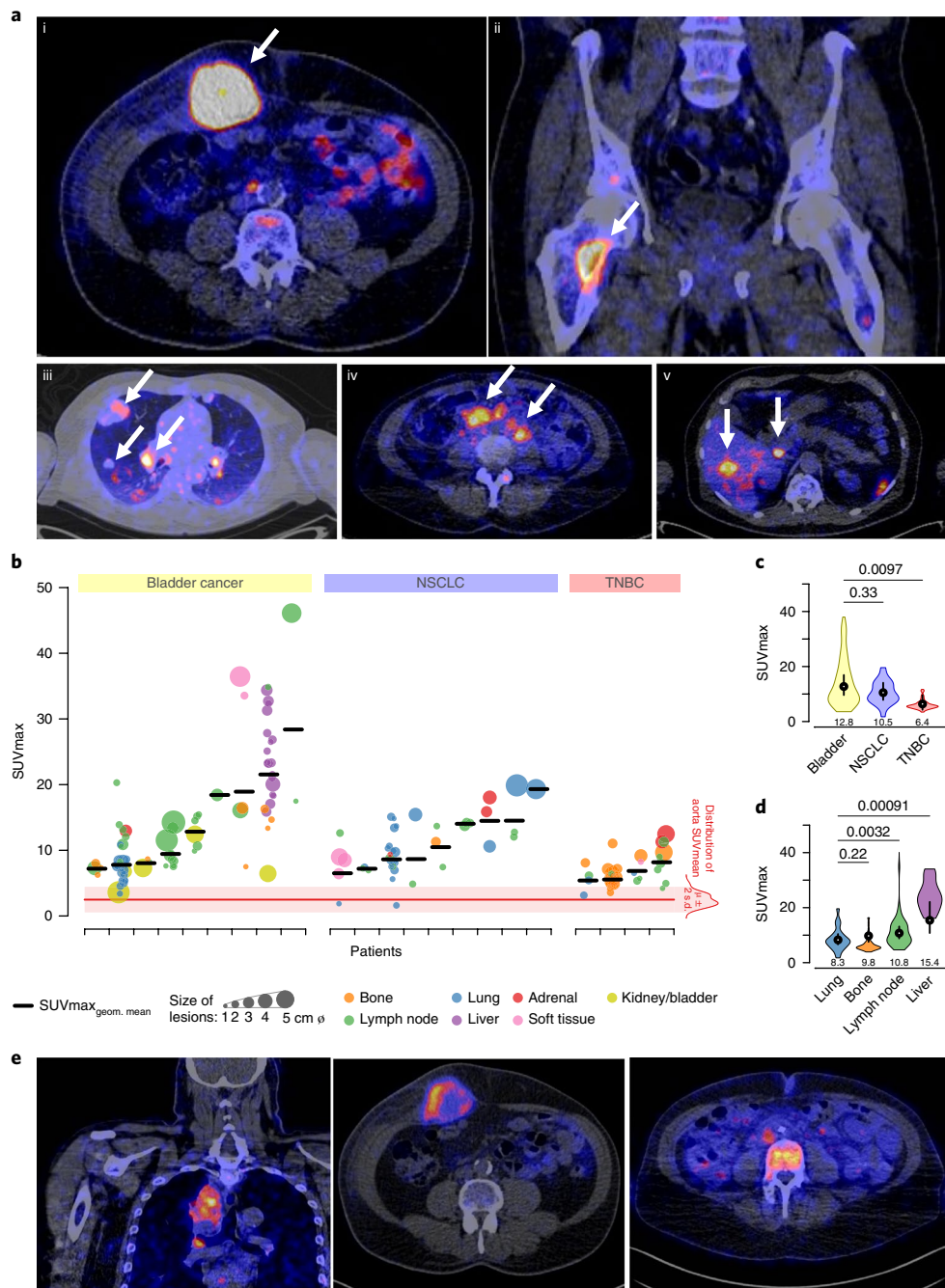
Fig. 2 | ⁸⁹Zr-atezolizumab tumor uptake. a, Examples of PET/CT images of four patients illustrating ⁸⁹Zr-atezolizumab tumor uptake in five different locations on day 7 postinjection (white arrows indicate tumor lesions; PET scans were performed once per patient and time point). Images (i) and (ii) are from the same patient, whereas images (iii), (iv), and (v) are from a separate patient each. **b**, Overview of ⁸⁹Zr-atezolizumab uptake as SUVmax at day 7 postinjection in 196 tumor lesions with a diameter >2 cm grouped per tumor type and ordered by increasing geometric mean SUVmax per patient, visualizing tumor size and site, and with blood pool background uptake as reference. Horizontal bars indicate geometric mean SUVmax per patient. **c**, Violin plot of actual distribution of SUVmax in lesions per site of lesion with bottom and top 1% of SUVmax values truncated (first, fiftieth, and ninety-ninth SUVmax percentile: 1.7, 7.9, 19.6 for lung; 3.9, 5.6, 16.4 for bone; 4.6, 9.7, 40.1 for lymph node; 16.1, 23.3, 34.1 for liver); black vertical lines are 95% CIs of geometric mean SUVmax, white dots within black lines and values below the violin plot are the actual geometric means, all based on a linear mixed regression model with two-sided Wald P values using Satterthwaite approximations to degrees of freedom under restricted maximum likelihood, shown above the graph; $n_{lung} = 44$ in ten patients, $n_{bone} = 62$ in nine patients, $n_{lymph\ node} = 54$ in 20 patients, $n_{liver} = 19$ in one patient. **d**, Violin plot of SUVmax in lesions per tumor type with bottom and top 1% of SUVmax values truncated (first, fiftieth, and ninety-ninth SUVmax percentile: 3.6, 10.9, 38.0 for bladder; 1.7, 9.7, 19.6 for NSCLC; 3.4, 5.6, 11.7 for TNBC); black vertical lines are 95% CIs of geometric mean SUVmax, white dots within black lines and values below the violin plot are the actual geometric means, all based on a linear mixed regression model with two-sided Wald P values using Satterthwaite approximations to degrees of freedom under restricted maximum likelihood, shown above the graph; $n_{bladder} = 85$ in nine patients, $n_{NSCLC} = 43$ in nine patients, $n_{TNBC} = 68$ in four patients. **e**, PET/CT images of lesions of three patients with heterogeneous intralesional ⁸⁹Zr-atezolizumab uptake on day 7 postinjection (PET scans were performed once per patient and time point). Mediastinal lesion of a NSCLC patient (SUVmax 19.9) (left), an abdominal wall metastases of a bladder cancer patient (SUVmax 36.4) (middle), and a bone metastasis of a TNBC patient (SUVmax 7.1) (right).
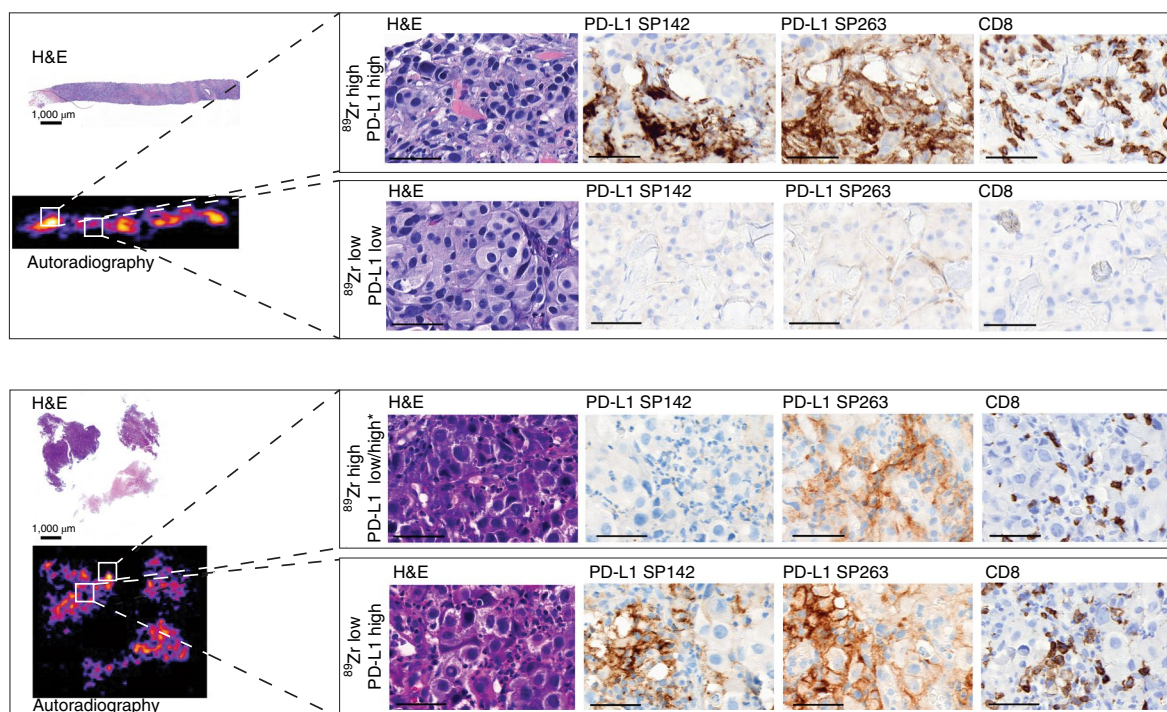
Fig. 3 | Autoradiography and IHC of postimaging tumor biopsy. PD-L1 IHC (SP142 and SP263), as well as CD8 IHC, of TNBC (upper panels). PD-L1 IHC (SP142 and SP263), as well as CD8 IHC, of bladder cancer biopsy samples (lower panels). Scale bars, 50 μm; autoradiography and all IHC were performed once per sample.

less uptake than bladder cancer (Fig. 2d). We also observed heterogeneous intratumor tracer distribution in several lesions of multiple patients (Fig. 2e). Autoradiography of two tumor samples showed heterogeneous tracer distribution, and PD-L1 as well as CD8 IHC showed heterogeneous staining, partly corresponding with regions of high tracer uptake (Fig. 3).

Given the high [89]Zr-atezolizumab tumor uptake, the known property of [89]Zr to remain in cells, and the potential role of atezolizumab internalization contributing to this signal, we determined the internalization of [89]Zr-atezolizumab in vitro in two tumor cell lines and in healthy volunteers' peripheral blood mononuclear cells (PBMCs) and in T cells. We observed high internalization rates in the tumor cell lines, and lower rates in human PBMCs and T cells (Supplementary Fig. 6a,b). Analyses of the PBMC fraction isolated from blood obtained from participating patients confirmed that only, 0.3% and 0.4% of the tracer dose on day 4 ($n = 2$) and 0.6% on day 7 ($n = 1$), respectively, was bound to and/or internalized by PBMCs. The lower internalization rates observed in PBMCs and T cells relative to the tumor lines are assumed to be primarily dependent on the lower PD-L1 expression of these cells.

To help explain why some patients respond to checkpoint inhibitors despite low or absent PD-L1 expression, we compared PD-L1 expression and immune phenotypes, on the basis of IHC and RNA sequencing of post-tracer biopsies with tumor tracer uptake.

As expected according to a prior study[17], the two IHC assays generated conflicting results in 8 of 19 samples (kappa 0.17, 95% CI −0.23 to 0.57; Supplementary Fig. 7a). [89]Zr-atezolizumab uptake of the biopsied tumor lesions increased with PD-L1 IHC scores for SP142 (Supplementary Fig. 7b) but not for SP263 (Supplementary Fig. 7c). Furthermore, tracer uptake did not differ between IHC-based immune phenotypes ($n = 16$; Supplementary Fig. 7d). The correlation of PD-L1 and T effector gene expression levels (CD8, granzyme B, interferon-γ, chemokine ligand 9, and combined T effector signature) with [89]Zr-atezolizumab ranged between 0.51 and 0.76 (Pearson), and 0.46 and 0.62 (Spearman), respectively

(Supplementary Fig. 7e–j). Similarly to the IHC results, this could partly explain the generally high tumor tracer uptake.

With a data cutoff date of 1 June 2018, seven patients were still in follow-up (median follow-up 21.9 months, range 16.7–25.1). Four of them were still on treatment, two were discontinued after two years of treatment and ongoing response, and one patient stopped atezolizumab, owing to side effects. Complete response (CR) was observed in three patients and partial response (PR) in four. Eleven patients showed stable disease (SD) as best response, and four patients progressed at the first CT evaluation (6 weeks; Supplementary Fig. 8a,b). The objective response rate was 56% for bladder cancer, 11% for NSCLC, and 25% for TNBC (Supplementary Table 3). Earlier reports have described objective response rates of 26% for unselected urothelial carcinoma patients[8,29], whereas we included only patients with bladder cancer, 21–23% for NSCLC patients[30,31], and 24% for PD-L1-positive TNBC patients[32]. The median progression-free survival (PFS) was 4.8 months (95% CI 2.7 to ∞) for all patients and 13.3 months (95% CI 4.1 to ∞) for those with SD, PR, or CR as best response.

At the patient level, [89]Zr-atezolizumab tumor uptake increased with increasing best tumor response category (Fig. 4a; log-linear trend) and was related to target lesion size change (Fig. 4b): patients with CR as best response had a 235% higher SUVmax (95% CI 98 to 467%; $P = 0.00021$) than patients who immediately progressed, and a twofold increase in geometric mean SUVmax per patient was associated with a best change in target lesion size from baseline of −35% on average (95% CI −61 to −9%; $P = 0.010$). Furthermore, the geometric mean SUVmax per patient was strongly related to PFS (14 events) and overall survival (OS; 11 events): those with below-median uptake were more likely to progress or die than those with above-median uptake (Fig. 4c). These relationships between [89]Zr-atezolizumab uptake and patient outcome did not change and remained significant following adjustment for tumor type and tumor load (Supplementary Table 4), and when analyzed continuously (Supplementary Table 5). In our patient population, which
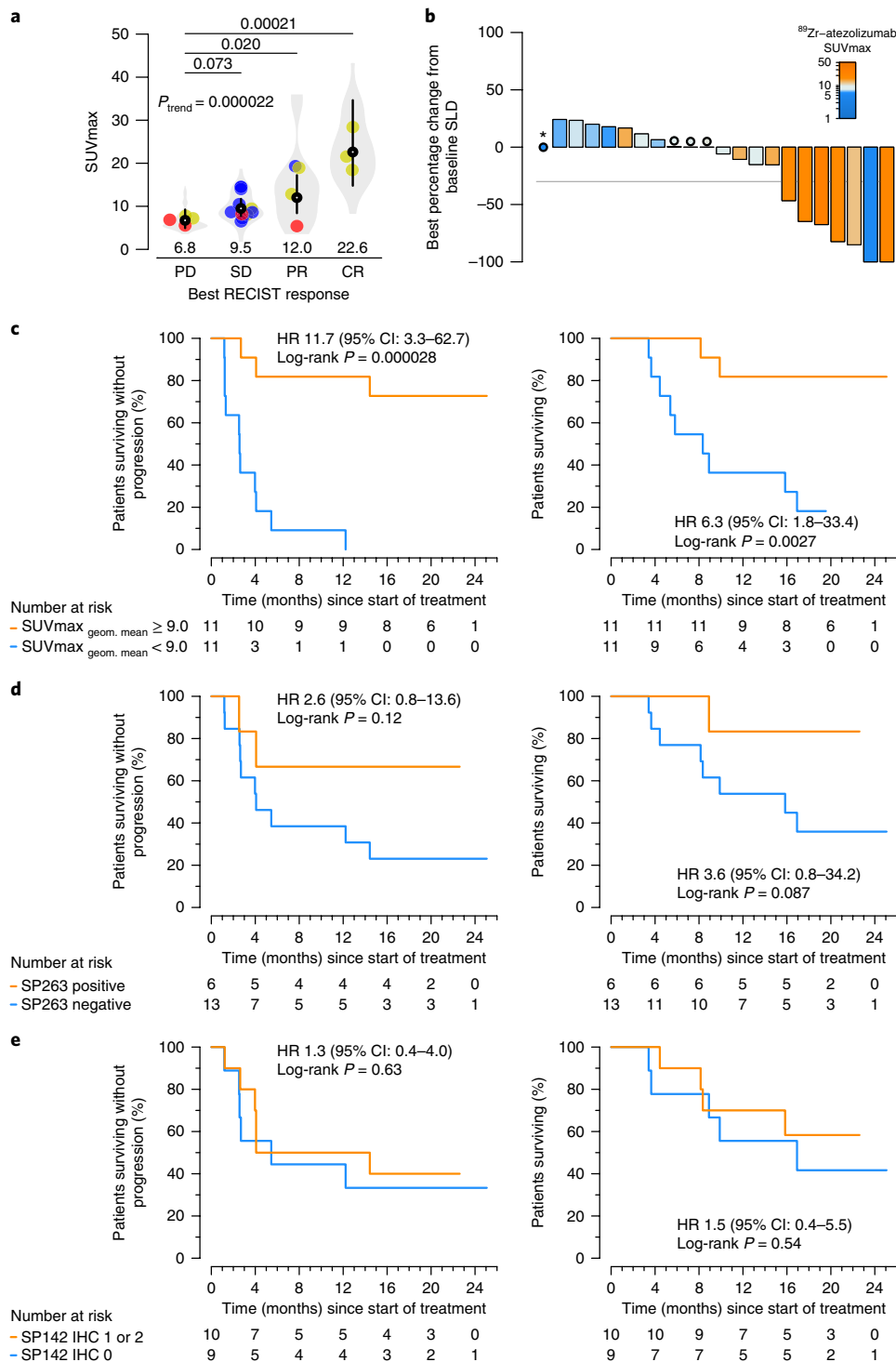
Fig. 4 | ^{89}Zr-atezolizumab tumor uptake as predictor for response. a, Relationship between ^{89}Zr-atezolizumab tumor uptake as geometric mean SUVmax and best RECIST response: gray violin plot areas show actual distribution of SUVmax at the metastasis level per best RECIST response category, with bottom and top 1% values truncated (first, fiftieth, and ninety-ninth SUVmax percentile: 3.5, 5.9, 14.7 for PD; 1.8, 90, 18.7 for SD; 3.5, 13.5, 36.0 for PR; 8.3, 23.2, 43.2 for CR); points show geometric mean uptake per patient, with colors indicating tumor type (red, TNBC; blue, NSCLC; yellow, bladder cancer); black vertical lines are 95% CIs of geometric mean SUVmax, and white dots within black lines and values below the violin plot are the actual geometric means; all are based on a linear mixed regression model with two-sided Wald *P* values using Satterthwaite approximations to degrees of freedom under restricted maximum likelihood, shown above the graph, and a two-sided *P* for trend based on a likelihood-ratio test under maximum likelihood; $n_{PD}=88$ in 4 patients, $n_{SD}=65$ in 11 patients, $n_{PR}=16$ in 4 patients, $n_{CR}=27$ in 3 patients). **b**, Waterfall plots depicting best percentage change from baseline SLD (measured on CT), with color scale indicating geometric mean SUVmax of tumor lesions per patient; circles show geometric SUVmax for patients with minimal change; asterisk indicates patient who was immediately progressive; no SLD change available. **c**, PFS and OS according to the geometric mean SUVmax per patient (orange, above-median geometric mean uptake; blue, below-median geometric mean uptake; $n=22$ patients; two-sided log-rank test). **d**, PFS and OS based on PD-L1 IHC (SP263: orange, IHC positive; blue, IHC negative; $n=19$ patients; two-sided log-rank test). **e**, PFS and OS based on PD-L1 IHC (SP142: orange, IHC positive; blue, IHC negative; $n=19$ patients; two-sided log-rank test).

lacked patients with a SP142 PD-L1 IHC score of 3, higher PD-L1 IHC expression was not related to better outcome (relationship with best tumor response: Fisher's exact $P = 0.71$ and 0.80, $n = 19$; relationship with best target lesion size change: t-test $P = 0.46$ and 0.42 for SP263 and SP142, respectively, $n = 18$; Fig. 4d,e shows relationship with survival). A patient's geometric mean [89]Zr-atezolizumab uptake discriminated effectively between patients with and without PR/CR as best tumor response, and between those with long and short time to progression or time to death (Supplementary Table 6). Both PD-L1 IHC assays, in contrast, showed moderate to poor discrimination for patient outcome (Supplementary Table 6).

At lesion level, [89]Zr-atezolizumab uptake was also related to change in size during treatment (Supplementary Fig. 9a–c). A multilevel linear mixed model—taking into account the repeated treatment response measurements during treatment per lesion and the clustering of lesions within patients (651 measurements from 107 metastases in 21 patients)—showed that the response of an individual lesion was strongly related to baseline SUVmax, with higher uptake indicating better response (interaction, (time since start treatment) × (SUVmax): $P = 5.1 \times 10^{-13}$; Supplementary Fig. 10).

In conclusion, in this first-in-human assessment of [89]Zr-atezolizumab, we show that the imaging signal corresponds to PD-L1 expression at sites of inflammation and in various normal lymphoid tissues. Furthermore, in our study (with a small patient population and no tumor biopsies that were immunohistochemically highly PD-L1 positive) tracer uptake appeared to be a strong predictor of response to atezolizumab treatment, including PFS and OS. Future clinical studies are needed to confirm our findings in a larger patient population, to comprehensively assess different [89]Zr-atezolizumab uptake features in combination with other clinical data to optimize therapy response prediction, and to evaluate whether [89]Zr-atezolizumab PET could also be used as a response predictor for treatment with other monoclonal antibodies targeting the PD-1/PD-L1 axis.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41591-018-0255-8

References

1. Balar, A. V. et al. Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: a single-arm, multicentre, phase 2 trial. *Lancet* **389**, 67–76 (2017).
2. Daud, A. I. et al. Programmed death-ligand 1 expression and response to the anti-programmed death 1 antibody pembrolizumab in melanoma. *J. Clin. Oncol.* **34**, 4102–4109 (2016).
3. Hegde, P. S., Karanikas, V. & Evers, S. The where, the when, and the how of immune monitoring for cancer immunotherapies in the era of checkpoint inhibition. *Clin. Cancer Res.* **22**, 1865–1874 (2016).
4. Rosenberg, J. E. et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. *Lancet* **387**, 1909–1920 (2016).
5. Chen, P. L. et al. Analysis of immune signatures in longitudinal tumor samples yields insight into biomarkers of response and mechanisms of resistance to immune checkpoint blockade. *Cancer Discov.* **6**, 827–837 (2016).
6. Ansell, S. M. et al. PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. *N. Engl. J. Med.* **372**, 311–319 (2015).
7. Dong, Z. Y. et al. Potential predictive value of TP53 and KRAS mutation status for response to PD-1 blockade immunotherapy in lung adenocarcinoma. *Clin. Cancer Res.* **23**, 3012–3024 (2016).
8. Fehrenbacher, L. et al. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. *Lancet* **387**, 1837–1846 (2016).
9. Kowanetz, M. et al. Tumor mutation load assessed by FoundationOne (FM1) is associated with improved efficacy of atezolizumab (atezo) in patients with advanced NSCLC. *Ann. Oncol.* **27**, 15–42 (2016).
10. Van Allen, E. M. et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* **350**, 207–211 (2015).
11. Kim, J. M. & Chen, D. S. Immune escape to PD-L1/PD-1 blockade: seven steps to success (or failure). *Ann. Oncol.* **27**, 1492–1504 (2016).
12. Heskamp, S. et al. Noninvasive imaging of tumor PD-L1 expression using radiolabeled anti-PD-L1 antibodies. *Cancer Res.* **75**, 2928–2936 (2015).
13. Chatterjee, S. et al. A humanized antibody for imaging immune checkpoint ligand PD-L1 expression in tumors. *Oncotarget* **7**, 10215–10227 (2016).
14. Rehman, J. A. et al. Quantitative and pathologist-read comparison of the heterogeneity of programmed death-ligand 1 (PD-L1) expression in non-small cell lung cancer. *Modern Pathol.* **30**, 340–349 (2017).
15. Smith, J. et al. Quantitative and qualitative characterization of two PD-L1 clones: SP263 and E1L3N. *Diagn. Pathol.* **11**, 44 (2016).
16. McLaughlin, J. et al. Quantitative assessment of the heterogeneity of PD-L1 expression in non-small-cell lung cancer. *JAMA Oncol.* **2**, 46–54 (2016).
17. Scheel, A. H. et al. Harmonized PD-L1 immunohistochemistry for pulmonary squamous-cell and adenocarcinomas. *Modern Pathol.* **29**, 1165–1172 (2016).
18. Madore, J. et al. PD-L1 expression in melanoma shows marked heterogeneity within and between patients: implications for anti-PD-1/PD-L1 clinical trials. *Pigm. Cell Melanoma Res.* **28**, 245–253 (2015).
19. Sanmamed, M. F. & Chen, L. Inducible expression of B7-H1 (PD-L1) and its selective role in tumor site immune modulation. *Cancer J.* **20**, 256–261 (2014).
20. Gniadek, T. J. et al. Heterogeneous expression of PD-L1 in pulmonary squamous cell carcinoma and adenocarcinoma: implications for assessment by small biopsy. *Modern Pathol.* **30**, 530–538 (2017).
21. McDermott, D. F. et al. Atezolizumab, an anti-programmed death-ligand 1 antibody, in metastatic renal cell carcinoma: long-term safety, clinical activity, and immune correlates from a phase Ia study. *J. Clin. Oncol.* **34**, 833–842 (2016).
22. Herbst, R. S. et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature* **515**, 563–567 (2014).
23. Dijkers, E. C. et al. Biodistribution of [89]Zr-trastuzumab and PET imaging of HER2-positive lesions in patients with metastatic breast cancer. *Clin. Pharmacol. Ther.* **87**, 586–592 (2010).
24. Gaykema, S. B. et al. [89]Zr-trastuzumab and [89]Zr-bevacizumab PET to evaluate the effect of the HSP90 inhibitor NVP-AUY922 in metastatic breast cancer patients. *Clin. Cancer Res.* **20**, 3945–3954 (2014).
25. Lamberts, T. E. et al. ImmunoPET with anti-mesothelin antibody in patients with pancreatic and ovarian cancer before anti-mesothelin antibody-drug conjugate treatment. *Clin. Cancer Res.* **22**, 1642–1652 (2016).
26. den Hollander, M. W. et al. TGF-β antibody uptake in recurrent high-grade glioma imaged with [89]Zr-fresolimumab PET. *J. Nucl. Med.* **56**, 1310–1314 (2015).
27. Ogembo, J. G. et al. SIRPα and FHOD1 are unique markers of littoral cells, a recently evolved major cell population of red pulp of human spleen. *J. Immunol.* **188**, 4496–4505 (2012).
28. Fullerton, J. N. & Gilroy, D. W. Resolution of inflammation: a new therapeutic frontier. *Nat. Rev. Drug Discov.* **15**, 551–567 (2016).
29. Johnson, D. B., Rioth, M. J. & Horn, L. Immune checkpoint inhibitors in NSCLC. *Curr. Treat. Options Oncol.* **15**, 658–669 (2014).
30. Powles, T. et al. MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer. *Nature* **515**, 558–562 (2014).
31. Petrylak, D. P. et al. Atezolizumab (MPDL3280A) monotherapy for patients with metastatic urothelial cancer: long-term outcomes from a phase 1 study. *JAMA Oncol.* **4**, 537–544 (2018).
32. Emens, L. et al. Inhibition of PD-L1 by MPDL3280A leads to clinical activity in patients with metastatic triple-negative breast cancer (TNBC). *Cancer Res.* **75**, abstr. 2859 (2015).

Author contributions

F.B., E.G.E.d.V., B.M.F., and A.d.C. designed the study. F.B., E.L.v.d.V., M.N.L.-d.H., A.J.-S., R.B., S.G.E., B.M.F., C.M., and A.d.C. developed the methodology. Acquisition of data was performed by F.B., C.M., T.C.K., E.L.v.d.V., I.C.K., S.F.O., C.P.S., T.J.N.H.,

Competing interests

Additional information

Methods

Patient population. Patients eligible for the study had histologically or cytologically documented locally advanced or metastatic bladder cancer, NSCLC, or TNBC. They were eligible for at least second-line systemic therapy, or, in the case of bladder cancer and NSCLC, showed disease progression during or within 6 months of completing platinum-based adjuvant/neoadjuvant chemotherapy. Other eligibility criteria included measurability according to RECIST 1.1[33], the presence of a tumor lesion from which a biopsy could safely be obtained, age ≥18 years, written informed consent, Eastern Cooperative Oncology Group performance status of 0–1, and adequate hematologic and end-organ function. Exclusion criteria were central nervous system disease, leptomeningeal disease, uncontrolled tumor-related pain, effusion/ascites, hypercalcemia, hypoalbuminemia, HIV infection, active tuberculosis, hepatitis B or C infections, current or recent severe infections, other substantial concomitant diseases including autoimmune diseases, recent treatment with systemic immunosuppressive or immunostimulatory medication, and prior treatment with CD137 agonists or immune-checkpoint inhibitors.

The imaging and treatment studies were performed in compliance with all relevant ethical regulations. Both studies were approved by the Medical Ethical Committee of the University Medical Center Groningen (UMCG) and the Central Committee on Research Involving Human Subjects, and were registered individually (ClinicalTrials.gov identifiers NCT02453984 and NCT02478099). All patients provided written informed consent.

Study design. This single-center, open-label imaging study was performed together with a companion atezolizumab treatment study at the UMCG, the Netherlands.

Patients received 10 mg unlabeled atezolizumab followed by 37 MBq (1 mCi) zirconium-89 (^{89}Zr)-atezolizumab (~1 mg antibody) intravenously (i.v.). Previous pharmacokinetic studies have shown that atezolizumab has dose-dependent kinetics[22,34]. Therefore, to reduce fast atezolizumab clearance, the additional unlabeled protein dose was added. In the first cohort of patients, ^{89}Zr-atezolizumab injection was followed by four PET scans at 1 h and at 2, 4, and 7 d postinjection (±1 d). In the subsequent cohort, imaging was performed by using the optimal schedule determined in the first cohort of patients. We considered the unlabeled antibody dose to be sufficient when the circulating amount of radioactivity on day 4 after ^{89}Zr-atezolizumab administration was comparable to that of other ^{89}Zr-monoclonal antibodies with well-known kinetics over time[23–26]. The optimal time point for PET scanning was determined by analyzing tumor tracer uptake and the available amount of radioactivity in the circulation.

After ^{89}Zr-atezolizumab administration, patients were observed for at least 1 h for infusion-related reactions. All PET scans were performed in combination with a low-dose computerized tomography (CT) scan for attenuation correction and anatomic reference with a Biograph mCT 64-slice PET/CT camera or a Biograph mCT 40-slice PET/CT camera (both Siemens). Within 7 d of the last PET scan, a tumor biopsy was obtained, after which 1,200 mg i.v. atezolizumab monotherapy was administered on a three-weekly schedule. Diagnostic CT scans were performed at baseline within 14 d before tracer injection and every 6 weeks (±3 d) after the start of atezolizumab treatment, or if clinically indicated.

^{89}Zr-atezolizumab production and in vivo stability. ^{89}Zr-atezolizumab was produced at the UMCG according to good manufacturing practice guidelines, as described earlier[35,36]. Quality control methods and manufacturing processes were validated before the start of clinical manufacturing. Stability testing was performed on both the conjugated intermediate and the ^{89}Zr-atezolizumab drug product. Release testing included size-exclusion high-performance liquid chromatography analysis for protein concentration and presence of aggregates, radiochemical purity testing, and an immunoreactivity assay to determine specific binding to PD-L1. ^{89}Zr-atezolizumab was produced with a specific activity of 37 MBq mg^{-1}, a radiochemical yield of >60%, and a radiochemical purity of >95%. The immunoreactivity of ^{89}Zr-atezolizumab was determined in a competitive binding assay with unlabeled atezolizumab, as described earlier[37]. The PD-L1 extracellular domain was used as a target. Unlabeled atezolizumab was added in a logarithmic concentration range of 0.5–2 mg ml^{-1}, and ^{89}Zr-atezolizumab was added with a fixed concentration of 1,000 ng ml^{-1}. The acceptance limit for immunoreactivity was >70%.

To determine the stability of ^{89}Zr-atezolizumab in vivo, a small subset of patients' serum samples obtained 2, 4, and 7 d postinjection were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis followed by phosphorimaging analysis. The fractions of bound ^{89}Zr-atezolizumab and free ^{89}Zr were calculated as a percentage of total radioactivity detected. ^{89}Zr-atezolizumab stored in saline at 2–8 °C served as a positive control.

^{89}Zr-atezolizumab PET. PET acquisition was dependent on the moment of scanning: 1 h after tracer injection a total body scan with up to 15 bed positions and 1.5 min per bed position was performed. At 2 and 4 d postinjection, the head to upper thigh was scanned in up to nine bed positions with 5 min per bed position and the legs in up to six bed positions with 2 min per bed position. For the first four subjects, on day 7, the head to upper thigh images were performed with up to nine bed positions with 10 min per bed position and the legs in up to six bed

positions with 4 min per bed position. To account for the lower count rate at day 7 and to increase image quality, subsequent patients were imaged with up to eight bed positions with 15 min per bed position. All PET images were reconstructed using the reconstruction algorithm recommended for multicenter ^{89}Zr-monoclonal antibody PET scan trials[38].

PET image analysis was performed with the Accurate tool for volume-of-interest (VOI)-based lesion and background analysis[39]. Spherical VOIs with predefined sizes were drawn in the thoracic aorta, subcutaneous tissue, liver, spleen, kidney, intestine, lung, brain, bone marrow and bone cortex, and muscle to assess ^{89}Zr-atezolizumab normal organ distribution. Tumor lesions were delineated manually or with the help of automated delineation algorithms based on the baseline diagnostic CT scan. To account for partial volume effects, for all statistical analyses only non-irradiated tumor lesions larger than 2.0 cm were selected.

SUVs were calculated by using the amount of injected activity, body weight, and the amount of radioactivity within a VOI. We report the SUVmax for tumor lesions and the SUVmean for normal organ tracer uptake.

In addition to VOI-based analysis, tracer uptake in non-malignant lymph nodes and the tonsils was compared qualitatively with surrounding tissue uptake and with tracer uptake in the healthy liver. The lymph nodes were defined as non-malignant on the basis of the diagnostic baseline CT scan by a dedicated radiologist, and tracer uptake was scored together with a dedicated nuclear medicine physician.

Safety assessment. Assessment of adverse events was performed at each clinical visit, and the National Cancer Institute Common Terminology Criteria for Adverse Events version 4.0 were used[40].

Pharmacokinetic assessments. Blood samples were collected for determination of tracer amount in the PBMC fraction and atezolizumab serum concentration before tracer injection, 1 h postinjection, and 2, 4, and 7 d postinjection.

Whole blood samples were collected in ethylenediaminetetraacetic acid blood tubes (BD) and fractionated into red blood compartment, plasma, and PBMC compartment through Ficoll-Paque Plus separation (GE Healthcare Life Sciences). The activity in different fractions was measured in a calibrated well-type LKB 1282 Compugamma system (LKB Wallac).

Atezolizumab serum concentration was measured by an indirect sandwich enzyme-linked immunosorbent assay (ELISA, performed by ICON Laboratory Services, Inc.). Values below the lower limit of quantification (60 ng ml^{-1}; $n = 1$) were substituted by half this lower limit (that is, 30 ng ml^{-1}) before further data analysis.

Atezolizumab internalization in vitro. Internalization of ^{89}Zr-atezolizumab was determined in the human lung mucoepidermoid pulmonary H292 and the bronchioalveolar H358 tumor cell lines (American Type Culture Collection, NCIH292 and NCIH358, respectively). ^{89}Zr-atezolizumab (50 ng) was added to 1×10^6 cells and incubated for 1 h on ice. Thereafter, cells were washed with ice-cold phosphate-buffered saline containing 1% human serum albumin (Sanquin), and the total amount of bound activity was measured in a gamma counter. Next, cells were incubated for 1, 2, and 4 h at 37 °C to allow internalization and at 4 °C to serve as a control. After incubation, cells were washed with acidic buffer, 0.05 M glycine (Merck) and 0.1 M NaCl (Merck) at pH 2.8, to remove the membrane-bound fraction. The amount of internalized activity was measured in a gamma counter. Internalized activity as a percentage of the total bound activity was determined. The mean standard deviation of three wells ($n = 3$) was calculated in Graphpad Prism 5.0.

The assay described above, with up to 2 h incubation time, was also performed with PBMCs (5×10^6 cells) freshly isolated from buffy coat pooled from healthy volunteers, as well as T cells (5×10^6 cells) isolated from PBMCs expanded with recombinant human interleukin-2 (100 U ml^{-1}, Novartis Pharma B.V.), anti-CD3 (clone UCHT1, 0.5 µg ml^{-1}, R&D Systems), and anti-CD28 (clone 37407, 2 µg ml^{-1}, R&D Systems) for 3 d. The percentage of internalization was calculated as the mean standard deviation of two wells ($n = 2$) in Graphpad Prism 5.0.

The internalization rate was estimated by saturating the cell surface with ^{89}Zr-atezolizumab at time zero and measuring the fraction of radioactivity that could no longer be displaced from the cell surface after incubation.

Tumor biopsies and normal lymph node and spleen tissue. PD-L1 expression was assessed centrally (HistoGeneX) in post-tracer tumor biopsies with the SP142 and the SP263 IHC assays (Ventana Medical Systems) according to the manufacturer's staining and scoring protocols. PD-L1 staining on tumor cells (TC) and on tumor-infiltrating immune cells (IC) was evaluated. With the SP142 assays, tumors were scored as negative (IC0 or TC0: staining on <1% of IC or TC, respectively; IHC score 0) or positive (IC1/2/3 or TC1/2/3: staining on ≥1% of IC or TC; IHC score 1/2/3 depending on the highest staining for either IC or TC). With the SP263 assay, tumors were considered PD-L1 positive when the membrane of ≥25% of tumor cells stained for PD-L1 at any intensity.

Infiltration and localization of CD8-positive T cells was assessed by using the clone C8/144B (Dako) to characterize histopathologically preexisting immunity of these tumors. Tumors were classified as immunological deserts if intratumoral

stroma areas or intraepithelial areas contained either no or few CD8[+] T cells, inflamed if ≥5% of the intraepithelial area was covered by CD8[+] T cells of intermediate density, or heterogeneous in case of any CD8 coverage in between.

Lymph node and spleen tissue, originating from spare healthy tissue not obtained from the study population, was also stained for PD-L1 and CD8 as described above, to correlate with [89]Zr-atezolizumab uptake in these tissues.

For revealing tracer distribution at the microscopic level, formalin-fixed tumor sections (10 μM) of post-tracer tumor biopsies of two patients were exposed to a phosphorimaging screen for 72 h and were then scanned with a Cyclone phosphorimager. Subsequent sections of the same tumor tissue were stained for hematoxylin and eosin, PD-L1, and CD8.

RNA from post-tracer tumor biopsies was isolated for gene expression analysis by TruSeq RNA Access RNA-seq (Q2Labsolutions).

PD-L1 IHC results were related to tumor SUVmax and clinical efficacy. Histopathological immune phenotypes and RNA expression levels (in reads per kilobase per million) of interferon-γ, chemokine ligand 9, granzyme B, CD8, and PD-L1, including combined T effector signature, were related to tumor SUVmax in an exploratory manner.

Data analysis. Results presented here had a clinical cutoff date of 1 June 2018.

Pharmacokinetic analyses and biodistribution. We evaluated the changes in atezolizumab concentration over time postinjection by using linear mixed effect models taking clustering within patients, and if applicable within tumors, into account as random effects, and with days postinjection as a fixed effect. Atezolizumab concentration was analyzed in serum and PET using the following parameters: the [89]Zr-atezolizumab SUVmean in healthy tissues, the [89]Zr-atezolizumab SUVmax in tumor lesions, and the tumor-to-background ratio ([89]Zr-atezolizumab tumor SUVmax divided by background SUVmean). We evaluated only tumor-to-background ratio changes in time for all tumors combined (compared with blood), and for lung and bone metastases (compared with healthy lung and bone marrow, respectively), because other tumor sites had too few data points for meaningful evaluation. Similarly, we evaluated only SUVmax changes over time according to tumor type for NSCLC and bladder cancer patients, because all TNBC patients underwent only two PET scans.

ELISA, SUVmax, and tumor-to-background values were natural-log-transformed before analysis to improve model fit, as these data were substantially right skewed. With back-transformation, the results of such models consequently yield estimates of geometric mean values of the response variable, and model coefficients can be expressed as the percentage change in the response variable per unit increase in the explanatory variable. We analyzed time postinjection as a categorical variable, thus resulting in estimates of the (geometric) mean atezolizumab concentration per day, as well as continuously, thus resulting in time–concentration curves. To assess potential non-linear relationships between time since tracer injection and the various tracer concentration measures, we evaluated time in days without transformation (linear), and also by adding a natural logarithmic (log-linear) or a quadratic term (parabolic curve) for time to the models. We then used the Akaike information criterion of each model to select the best representation of the data. The relationship between tumor SUVmax values at day 4 and day 7 was also assessed with a mixed effect model, and by estimating the Pearson's correlation coefficient extended to clustered data, by following natural-log-transformation of both variables to obtain approximate normal distributions. We evaluated the correlation between ELISA-derived atezolizumab concentration (ng ml[−1], corrected for hematocrit) and PET-derived atezolizumab concentration (ng ml[−1]) in a similar way. The PET-derived atezolizumab concentration was interpolated by multiplying the SUVmean of the aorta with the injected total antibody mass averaged for body weight.

Tumor tracer uptake across and within patients. These analyses were based solely on day 7 postinjection tracer uptake measurements, because the SUVmax of day 4 and day 7 were highly correlated. We first estimated the geometric mean SUVmax across all tumors, taking between-patient heterogeneity into account by using an intercept-only linear mixed effect model with patient as the random effect and the natural-log-transformed SUVmax as the response variable. Then, to provide insight in the heterogeneity of [89]Zr-atezolizumab uptake, we plotted the SUVmax of each tumor lesion, grouped per tumor type and ordered by increasing geometric mean SUVmax per patient, also visualizing tumor size and site, and with blood pool background uptake (SUVmean) as a reference. To further assess the within-patient heterogeneity in tracer uptake for patients with at least one tumor site, we assessed the relative difference per patient between the tumor sites with the highest versus lowest uptake, as well as the coefficient of variation (that is, the standard deviation divided by the mean SUVmax of all tumor sites per patient after natural-log-transformation, expressed as a percentage). We used linear mixed effect models with patient as a random effect to evaluate whether the geometric mean SUVmax differed according to tumor organ site and according to the primary tumor type, separately.

Tumor tracer uptake compared with biopsy-derived molecular analyses. First, we evaluated the agreement between the SP142 and SP263 IHC analysis of PD-L1

expression in tumor biopsies by constructing confusion matrices (overall and according to tumor type) and by estimating kappa statistics.

Next, we assessed the associations among SP142, SP263, and immune phenotype levels with day 7 postinjection [89]Zr-atezolizumab SUVmax of the same lesions, by using independent sample *t*-tests and ordinary linear regression to test for a trend in increasing tracer uptake with increasing SP142 IHC levels (that is, fitted as a continuous variable coded 0, 1, or 2). These analyses were conducted after natural-log-transformation of SUVmax, thus yielding estimates of geometric means following back-transformation of the results.

We also compared the relationship between RNA-sequencing-derived estimates of CD8, granzyme B, interferon-γ, and chemokine ligand 9 gene expression levels, as well as a combined 'T effector signature' (that is, the per-biopsy average of the previous individual gene expression levels) with day 7 postinjection [89]Zr-atezolizumab SUVmax. For this, RNA-sequencing data (in reads per kilobase per million) and SUVmax were natural-log-transformed to achieve approximate normality before analysis by linear regression and before estimating Pearson's and Spearman's correlation coefficients (the latter to also include a more robust correlation measure in view of the small number of data points (<30) for these analyses).

Clinical outcome. Clinical outcome was evaluated in several ways. At the patient level, we evaluated the best-achieved RECIST response category, the best percentage change in the sum of largest diameters (SLD) of target lesions, and the PFS and OS. We first used general descriptive methods to summarize clinical outcomes, including a swimmers plot, the median time of follow-up (assessed in patients still without progression at data cutoff), and the median time to progression. Next, we assessed the relationship between [89]Zr-atezolizumab SUVmax at day 7 and clinical outcome. For these analyses, SUVmax was first natural-log-transformed and the results are thus reported as geometric means, or percentage changes in the average SUVmax. We specifically decided to summarize [89]Zr-atezolizumab uptake per patient as geometric mean SUVmax beforehand, and we refrained from evaluating different [89]Zr-atezolizumab uptake features per patient to prevent overoptimistic results due to evaluating many features in a small dataset.

The relationship between the best RECIST response categories and tumor lesion SUVmax was assessed with a linear mixed model accounting for within-patient clustering by a random intercept, and by evaluating best RECIST response both categorically and continuously (that is, as a 0, 1, 2, 3 variable). When the data were modeled continuously to test for a trend, we evaluated not only a linear fit but also the addition of a natural logarithmic or a quadratic term, and we used the Akaike information criterion to select the best representation of the data. The relationship between per patient geometric mean SUVmax and change over time in SLD was visualized with a spaghetti plot, whereas tracer uptake in relation to best percentage change in SLD of target lesions was visualized using a waterfall plot and further analyzed by linear regression and estimating the Pearson's correlation coefficient. The relationship between per patient geometric mean SUVmax and PFS and OS was explored by Kaplan–Meier survival plots, binning patients in a below-median and above-median geometric mean SUVmax group, and testing for survival differences using the log-rank test. We further quantified the relationship between these patient groups with high and low tracer uptake by fitting Cox regression models (with Firth's penalization to account for small sample bias), thus yielding hazard ratios for progressive disease and/or death. We decided a priori to bin patients at the median to achieve two equally sized groups to compare, and we refrained from exploring other/optimal thresholds in tracer uptake, because the small dataset precluded meaningful threshold finding. In addition to binning, we also evaluated the relationship between PFS and OS and geometric mean SUVmax per patient continuously, again using Firth's penalized Cox regression models, and expressing the results per standard deviation change in SUVmax (assuming a log-linear relationship, because the dataset was deemed too small to properly evaluate departure from linearity).

To evaluate to what extent the observed relationships between tracer uptake and clinical outcome could be explained by differences between patients in tumor types and tumor load, we also adjusted for these variables by including them in the various regression and mixed regression models (tumor type categorically, tumor load—the number of tumor lesions with SUVmax data per patient—linearly). Because the number of events for the survival analyses was rather small to estimate robust effects of these potential confounders together with tracer uptake in one model, we alternatively also adjusted the relationship between tracer uptake and PFS and OS by means of an inverse probability weighting (IPW) procedure based on propensity scores[41], which may be statistically more efficient in small datasets. In this two-step procedure, first the predicted probability (that is, the propensity score) of belonging to the below- or above-median geometric mean SUVmax group was estimated by a logistic regression model containing the potential confounders, thus resulting in a weight for each patient by taking the reciprocal of this probability (that is, the IPW); in the second step, a regular Cox regression model for PFS or OS was used with only tracer uptake group as the explanatory variable, while weighting patients by their IPW, thereby adjusting the association between tracer uptake and PFS or OS for the potential confounders. We performed these survival analyses in the entire study group and separately in a subgroup while

excluding the breast cancer cases, because these cases all belonged to the below-median SUVmax group, thus prohibiting adequate adjustment for this tumor type due to colinearity. We checked whether IPW adjustment was adequate by assessing the C index of the propensity model refitted using its own weights (values close to 0.5 indicating perfect adjustment), and by comparing the confounder distribution between the groups with low and high tracer uptake following IPW.

We evaluated the ability of per patient geometric mean ^{89}Zr-atezolizumab SUVmax to discriminate between patients with complete or partial response versus stable or progressive disease as best response by estimating the area under the receiver operating characteristics curve. The C index was used to assess the ability to discriminate between patients with short and long (progression-free) survival. An area under the curve or C index of 1.0 indicates perfect discrimination, and 0.5 indicates a worthless test. For these analyses, the per patient geometric mean ^{89}Zr-atezolizumab SUVmax was evaluated continuously instead of being binned at the median.

Similarly to the above, we also assessed the relationship between IHC-based PD-L1 expression by SP263 and SP142 antibodies and clinical outcome. As these tests were analyzed dichotomously (that is, positive or negative), their relationship with best RECIST response category was tested by the Fisher's exact test, and their relationship was tested with best percentage change in SLD by independent sample t-tests.

At the lesion level, we used spaghetti plots to visualize the relationship between day 7 ^{89}Zr-atezolizumab SUVmax and tumor size change during treatment. We then modeled the tumor size change during treatment as a function of baseline SUVmax, using a linear mixed effect model with per-lesion percentage change in size at the various follow-up moments as response variable, time since start of treatment and SUVmax as explanatory variables, and patient- and lesion-level random intercepts. Both time and SUVmax were natural-log-transformed in these analyses, and a fixed effect intercept was omitted to force the regression lines through the origin (that is a percentage change in size of 0 at time 0). We used an interaction term between time and SUVmax to test whether the change in lesion size during treatment depended on the baseline SUVmax.

Statistical inference. All reported P values are two sided with a threshold for statistical significance of 5%, and estimates and measures of association are reported with 95% CIs. We did not account for multiple testing, owing to the exploratory nature of this study, and further studies are thus needed for confirmation of the results. Analyses were performed using R software version 3.2.1 for Macintosh, specifically using the lmer function (libraries lme4 version 1.1-11 and lmerTest version 2.0-20) for mixed effect models; the Cohen.kappa function (library psych version 1.4.8) for estimating kappa statistics; the roc function (library pROC version 1.8) for estimating areas under curves; the survfit, survdiff, and coxph functions (library survival version 2.38-3) for Kaplan–Meier, log-rank, Cox regression analyses and estimating C indexes; and the coxphf function (library coxphf version 1.11) for Cox regression analyses with Firth's penalization.

For regression analyses, we used Wald-based tests; for instance, to compare different levels of a categorical variable (for example, specific tumor organ sites and primary tumor types), or likelihood-ratio-based tests to globally test the contribution of a categorical variable or test for interactions. Wald-based P values of mixed effect model coefficients and 95% CIs were based on Satterthwaite approximations to degrees of freedom under restricted maximum likelihood. Likelihood-ratio-based P values of mixed effect models were obtained under maximum likelihood.

P values and 95% confidence intervals for the IPW-adjusted survival analyses were based on 4,000-fold bootstrap resampling, repeating the full two-step analysis approach within each bootstrap, and using the bootstrap percentile method for inference.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The RNA-sequencing dataset presented in this manuscript is available through GEO (series accession GSE115594). The data are annotated with a short summary and a description of the study design and can be freely downloaded via the GEO website (https://www.ncbi.nlm.nih.gov/geo/). Clinical details of the cases and laboratory data, restricted to non-identifying data owing to privacy concerns, can be requested by e-mail from the corresponding author, who will handle all requests.

References

33. Eisenhauer, E. A. et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur. J. Cancer* **45**, 228–247 (2009).
34. Lamberts, L. E. et al. Antibody positron emission tomography imaging in anticancer drug development. *J. Clin. Oncol.* **33**, 1491–1504 (2015).
35. Verel, I. et al. ^{89}Zr immuno-PET: comprehensive procedures for the production of ^{89}Zr-labeled monoclonal antibodies. *J. Nucl. Med.* **44**, 1271–1281 (2003).
36. Terwisscha van Scheltinga, A. G. et al. ImmunoPET and biodistribution with human epidermal growth factor receptor 3 targeting antibody ^{89}Zr-RG7116. *mABs* **6**, 1051–1058 (2014).
37. Oude Munnink, T. H. et al. PET with the ^{89}Zr-labeled transforming growth factor-beta antibody fresolimumab in tumor models. *J. Nucl. Med.* **52**, 2001–2008 (2011).
38. Makris, N. E. et al. Multicenter harmonization of ^{89}Zr PET/CT performance. *J. Nucl. Med.* **55**, 264–267 (2014).
39. Frings, V. et al. Repeatability of metabolically active tumor volume measurements with FDG PET/CT in advanced gastrointestinal malignancies: a multicenter study. *Radiology* **273**, 539–548 (2014).
40. U.S. Department of Health And Human Services. Common terminology criteria for adverse events (CTCAE) version 4.0 (NIH publication no. 09-7473) (National Cancer Institute, Bethesda, MD, USA, 2009).
41. Austin, P. C. & Stuart, E. A. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat. Med.* **34**, 3661–3679 (2015).

nature research

Corresponding author(s): E.G.E. de Vries

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	OptiQuant version 3.00 (SDS analysis); Image J version 1.50i (autoradiography); Caseviewer version 2.1 (IHC); raw BCL data was demultiplexed using Illumina's bcl2fastq v2.20 (https://support.illumina.com/downloads/bcl2fastq-conversion-software-v2-20.html) and adapters were clipped and reads were quality trimmed using fastq-mcf v1.05 (RNA-seq); Siemens mCT PET/CT camera software versions VG50A/VG51A/VG60A and Siemens CT camera software VA48A and Syngo.via version 10.01.03 (Imaging software); OpenClinica version 3.6 (case record form)
Data analysis	R software version 3.2.1 for Mac; Accurate tool for PET data analysis; Graphpad Prism 5.0

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The RNA sequencing data set presented in this manuscript is available through NCBI GEO (series accession number GSE115594). All other data supporting the findings of this study are available from the corresponding author upon request.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences

Study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Imaging feasibility trial with no pre-defined sample size, therefore no sample size calculations were performed. A total of 25 patients were enrolled. Three patients dropped out prematurely, therefore, 22 patients were evaluable for final analysis.
Data exclusions	Three patients were discontinued prematurely from the trial due to early disease progression not allowing completion of the imaging series and start of treatment.
Replication	Patients were injected with 89Zr-atezolizumab only once and underwent a PET scan on days 0, 2, 4 and 7 or 4 and 7, respectively. PET scans could not be repeated on the same time point in the same patient. All IHC was performed once per sample per staining, which was considered sufficient as IHC assays are validated assays which have been performed together with positive and negative controles. Autoradiography was performed only once per sample as patients were injected with 89Zr-atezolizumab and biopsied only once and scan procedure lasted 72 hours not allowing replication due to decay of 89Zr. SDS-PAGE was performed once per sample, as patients were injectd with 89Zr-atezolizumab only once and blood was only collected once per time point and used for several analysis. Assessment of 89Zr-atezolizumab activity in plasma/RBC and PBMCs of patients blood was performed once , as patients were injectd with 89Zr-atezolizumab only once and blood was only collected once per time point and used for several analysis. RNA sequencing was performed once per sample, which was considered sufficient as the test runs with internal controls. 89Zr-atezolizumab internalization in H292 and H358 cells was assessed once in 3 technical replicates. 89Zr-atezolizumab internalization in PBMCs and T cells of healthy volunteers was assessed once in 2 technical replicates.
Randomization	Randomization was not performed in this trial as it was a first-in-human feasibility trial and all patients received treatment with atezolizumab after molecular PET imaging.
Blinding	Blinding was not necessary in this trial, as all patients received treatment with atezolizumab after molecular PET imaging.

Materials & experimental systems

Policy information about availability of materials

n/a	Involved in the study
☒	☐ Unique materials
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Research animals
☐	☒ Human research participants

Antibodies

Antibodies used	- Therapeutic antibody: Atezolizumab, monoclonal antibody against PD-L1; provided by Hoffmann-La Roche/Genentech and approved for therapeutic use (detailed information can be found on the manufacturers website: https://www.roche.com/products.htm --> TECENTRIQ (atezolizumab); website accessed on August 8, 2018) - IHC, primary antibodies: Anti-PD-L1: Clone name: SP142, Supplier: Ventana Medical Systems, catalog number: 740-4859, lot-number: N/A, dilution: prediluted (dispenser); secondary antibody: OptiView DAB Detection Kit (Ventana 760-700) and OptiView Amplification Kit

(Ventana 860-099).
Anti-PD-L1: Clone name: SP263, Supplier: Ventana Medical Systems, catalog number: 790-4905, lot-number: N/A, dilution: prediluted (dispenser); secondary antibody: Optiview DAB IHC Detection (Ventana - 760-700).
Anti-CD8: Clone name: C8/144B, Supplier: Dako, catalog number: M7103, lot-number: various, dilution: lot specific; secondary antibody: UltraView Universal DAB detection kit (Ventana - 760-500)

Validation	-Therapeutic antibody: Atezolizumab, monoclonal antibody against PD-L1; provided by Hoffmann-La Roche/Genentech and approved for therapeutic use (detailed information can be found on the manufacturers website: https://www.roche.com/products.htm --> TECENTRIQ (atezolizumab); website accessed on August 8, 2018) - IHC, primary antibodies: Detailed information on validation can be found on the suppliers website (all websites accessed on August 8, 2018): SP142 --> http://www.ventana.com/product/1827?type=2357 SP263 --> http://www.ventana.com/product/1815?type=2324 C8/144B --> https://www.agilent.com/en/products/immunohistochemistry/antibodies-controls/primary-antibodies/cd8#productdetails - IHC, secondary antibodies: Detailed information on validation can be found on the suppliers website (all websites accessed on August 8, 2018): OptiView DAB Detection Kit (Ventana 760-700) --> http://www.ventana.com/product/1574?type=2037 OptiView Amplification Kit (Ventana 860-099) --> http://www.ventana.com/product/1714?type=2170 UltraView Universal DAB detection kit (Ventana - 760-500) --> http://www.ventana.com/product/1414?type=1791

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	human lung mucoepidermoid pulmonary H292 (American Type Culture Collection NCIH292) bronchioalveolar H358 tumor cell line (American Type Culture Collection NCIH358)
Authentication	The used cell lines were STR profilled before use.
Mycoplasma contamination	Cell lines were tested mycoplasma free.
Commonly misidentified lines (See ICLAC register)	No cell lines used in this study are commonly misidentified.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Patients eligible for the study had histologically or cytologically documented locally advanced or metastatic bladder cancer, NSCLC or TNBC. They were eligible for at least second-line systemic therapy, or – in case of bladder cancer and NSCLC, showed disease progression during or within 6 months of completing platinum-based adjuvant/neoadjuvant chemotherapy. Other eligibility criteria included measurability according to RECIST 1.1, presence of a tumor lesion from which a biopsy could safely be obtained, age ≥ 18 years, written informed consent, Eastern Cooperative Oncology Group performance status of 0-1 and adequate hematologic and end organ function. Exclusion criteria were central nervous system disease, leptomeningeal disease, uncontrolled tumor-related pain, effusion/ascites, hypercalcemia, hypoalbuminemia, HIV infection, active tuberculosis, hepatitis B or C infections, current or recent severe infections, other significant concomitant diseases including autoimmune diseases, recent treatment with systemic immunosuppressive or immunostimulatory medication and prior treatment with CD137 agonists or immune checkpoint inhibitors.

Method-specific reporting

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ Magnetic resonance imaging