Article in Press

High-throughput spheroid-based assay for functional breast cancer precision medicine facilitated by deep learning

Ben Haspels, Maarten W. Paul, Jayant Jagessar Tewari, Mieke Bavelaar, Zofia M. Komar, Jos Jonkers, Roland Kanaar, Dik C. van Gent & Maayke M. P. Kuijten

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

High-throughput spheroid-based assay for functional breast cancer precision medicine facilitated by deep learning

Ben Haspels[1,2], Maarten W. Paul[1,2,3], Jayant Jagessar Tewari[1], Mieke Bavelaar[1], Zofia M. Komar[1], Jos Jonkers[2,4], Roland Kanaar[1,2], Dik C. van Gent[1,2] and Maayke M.P. Kuijten[1,2]*

[1]Department of Molecular Genetics, Erasmus MC Cancer Institute, Erasmus University Medical Center, Rotterdam, the Netherlands.

[2]Oncode Institute, Erasmus University Medical Center, Rotterdam, the Netherlands.

[3]Department of Cell and Chemical Biology, Leiden University Medical Centre, Leiden, The Netherlands.

[4]Division of Molecular Pathology, The Netherlands Cancer Institute, Amsterdam, the Netherlands.

*Corresponding author: Maayke Kuijten
m.kuijten@erasmusmc.nl

Key words: breast cancer, precision medicine, treatment response, high-throughput, tumor spheroids, deep learning

Abstract

Background

Despite rapid advances in treatment, breast cancer remains the leading cause of cancer mortality in women, with triple negative breast cancers having a particularly poor prognosis. Some tumors have (epi)genetic alterations causing homologous recombination deficiency, providing opportunities for targeted therapeutics including poly (ADP-ribose) polymerase inhibitors. However, the effects of targeted treatments are variable; therefore, functional assays are needed to predict the best personalized treatment options.

Methods

We developed a high-throughput spheroid-based assay using patient-derived breast cancer xenograft models sensitive and resistant to cisplatin. Methods were developed for automatic spheroid segmentation using deep learning to measure response of spheroids to treatment with cisplatin, olaparib and radiotherapy. We developed a method to distinguish between sensitive and resistant tumors based on predicting the percentage of responding and non-responding spheroids.

Results

Here we show that differences in treatment response between cisplatin-sensitive and resistant tumors faithfully correspond with the expected *in vivo* responses. The assay is able to discriminate between olaparib-sensitive and resistant tumors based on predicting the percentage of responding and non-responding spheroids.

Conclusions

We demonstrate that this assay, guided by automatic spheroid segmentation using deep learning, may report on the tumor's sensitivity to therapies with the potential to be applied to functional precision oncology for breast cancer.

Plain language summary

Many women die due to breast cancer, especially if tumors are more difficult to treat or spread more easily to other organs. Some of these tumors can be successfully treated with drugs because of faults in their DNA repair. To choose the best treatment for the patient, it would be helpful to know which tumors are vulnerable to treatment. We therefore developed a test. We grew small spheres from tumor material (spheroids) and treated them with different drugs used in cancer therapies. To measure the response, we took pictures of the spheroids and we taught a computer program to recognize the spheroids and measure their size. We tested our assay with patient tumors grown on mice (sensitive and resistant to a drug). We demonstrate that this assay may report on the tumor's sensitivity to therapies and may be used in the future to help choose the best treatment for breast cancer.

Introduction

Despite the large number of treatment options, breast cancer is the most frequent cause of cancer death amongst women worldwide[1,2]. Although many patients are cured by surgical removal of the primary tumor, often in combination with radiotherapy, chemotherapy and/or anti-hormonal treatment, approximately 20% of the patients develop metastasized disease, which is much more difficult to treat. Especially patients with tumors that lack expression of Estrogen and Progesterone Receptors (ER or PR) and Human EGF Receptor 2 (HER2), the so-called triple negative breast cancers (TNBC), have a dismal prognosis[3]. Generally, these patients are treated with chemotherapy. Additionally, some cancers have acquired mutations in DNA repair

genes, particularly those involved in homologous recombination repair (HRR), providing opportunities for targeted therapeutics. However, development of therapy resistance necessitates chemotherapy instead of targeted therapy also in this group of patients. Moreover, not all patients benefit from treatments, especially in later lines of chemotherapy. Hence, personalized treatment regimens are needed to optimize therapy outcome and reduce adverse effects.

Precision oncology needs to ensure that a targeted therapy is only given to patients whose tumors harbor the characteristics that predict beneficial effects from this treatment. In TNBC, germline mutations in the *BReast CAncer gene 1 DNA repair associated* (*BRCA1*) and *BRCA2* genes are more common than in receptor positive subtypes, up to 10-20%[4]. Tumors with these mutations are generally homologous recombination deficient (HRD) causing sensitivity to Poly-[ADP-ribose]-polymerase inhibitors (PARPi)[5,6] and DNA damaging agents such as platinum-based chemotherapies. Therefore, accurate detection of HRD is important for treatment decision. Clinically-available sequencing-based tests to determine HRD status including the MYRIAD/ MyChoice[7,8], the BRCA1/2-like classifier[9], HRDetect[10] and Classifier of Homologous Recombination Deficiency (CHORD)[11] tests have their limitations[12,13]. Functional assays could complement sequencing approaches to address the limitations posed by genomics-based precision oncology. For instance, the RECAP test determines the HRD status of tumors based on the number of nuclear RAD51 foci upon X-ray treatment of tumor specimens[14,15,16,17]. However, the RECAP test only determines HRD status and cannot directly give information about expected treatment response. Functional assays that report on treatment response require personalized tumor models as patient avatars.

In vivo, tumors have complex cell-cell communication and signaling pathways due to their 3D structure and heterogeneity, which are lacking in cell monolayers, and therefore may respond

differently to DNA damaging agents[18]. Hence, 3D cell models, such as spheroids, organoids or tissue explants would better mimic *in vivo* conditions and may serve as patient avatars[12]. Spheroids consist of a spherical aggregate of tumor cells, formed inside a scaffold such as a hydrogel, inside a drop (hanging drop method) or using culture plates with low-adhesive surface, minimizing cell-surface binding and increasing cell-cell interactions, and thus stimulating cell aggregation[19]. Protocols to produce spheroids on a large scale have been developed[20,21,22], which makes these cell models suitable for high-throughput drug screening. For example, spheroids have been used in several preclinical studies to model treatment response to chemotherapy[12,23,24,25]. Moreover, spheroid culturing platforms are especially useful for readouts that require results within a few weeks as spheroid formation is fast, within 5-7 days after cell seeding.

To fully evaluate the heterogeneity of a tumor, high-throughput assays may be helpful in assessing both the average tumor response and heterogeneity in response. Tumor volume estimation is one of the most frequently used parameters to determine treatment response[26] and therefore spheroid size has been used to assess this[23]. To determine changes in spheroid size, various image analysis approaches can be used to automate spheroid segmentation and size measurements. Traditional methods rely on signal thresholds and other image characteristics such as circularity[27]. More recently, deep learning approaches, such as using convolutional Neuronal Networks (CNN) have successfully been applied for spheroid segmentation[21]. Alternatively, Segment Anything Model (SAM)[28], which uses a transformer-based computer vision model can be used for versatile prompt-guided object segmentation. The pre-trained model can be further fine-tuned to automatically segment specific objects from images, such as spheroids in bright-field microscopy images. Therefore high-throughput spheroid production in combination with these advanced deep learning approaches is an attractive approach for drug screening[22,29].

Here, we show the development of a high-throughput spheroid assay using patient-derived xenograft (PDX) breast cancer mouse models, reproducing the expected *in vivo* sensitivity of the models. By implementing deep-learning-based automatic spheroid segmentation, this approach has potential applications in screening primary patient material to guide personalized treatment strategies.

Methods

Breast PDX tumor models

Breast cancer PDX models used in this study included a HR deficient triple-negative breast tumor (c.2210delC/2329delC) sensitive to cisplatin and a cisplatin-resistant model (acquired resistant by repeated cisplatin treatment)[30] and tumor extraction from these mouse models was approved by the Central Committee on Animal Experiments (CCE) and by the Institute for Animal Welfare of the Netherlands Cancer Institute. Tumor fragments from PDX mouse models[30] were transplanted into the 4th mammary fat pad of 6-10 week-old female mice (NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ; Jackson Laboratory) under anesthesia with isoflurane (3-4% induction and 2-3% maintenance) and analgesia with Carprofen (in drinking water, a day before surgery and 3 days after, 0.067 mg/ml). The perpendicular tumor diameters of formed tumors were measured twice per week using a calliper, and tumor volume was calculated using vol(mm3) = 0.5(length × width2). Maximum permitted tumor volumes were 1000 mm3. Mice were killed by CO_2 asphyxiation to collect tumors post-euthanasia. Mice were kept in individually ventilated cages, with a 12-hour light/dark cycle (7am:7 pm) at 21 °C (-/+3°C) and

humidity of 55% (+/-10%). Food and water were provided ad libitum. The maximal permitted disease endpoints were not exceeded in any of the experiments. All experiments were performed under ethical approval AVD3010020172464 and work protocol 9.1.10374.

RECAP assay

The HR status of used PDX tumors was validated using the RECAP assay[15,17]. Briefly, the amount of RAD51 foci was determined 2 hours post-radiation with 5 Gy. As HR would only be functional in S/G2 cells, we used geminin staining and an additional 53BP1 staining was performed to verify proper induction of DSBs following treatment. PDX tumors were classified as HRP or HRD when more than 50% or less than 20% of geminin+ cells displayed more than five RAD51 foci, respectively.

Cell isolation from breast PDX tumors

For cell isolation from PDX material, the tumors were dissected into small pieces and collected in Dulbecco's Phosphate Buffered Saline (PBS, D8537, Sigma-Aldrich). The collected material was then centrifuged at 200xg for 5 minutes at room temperature. The material was further processed by re-suspension in pre-warmed (37°C) $1.6*10^4$ units of collagenase (C9407, Sigma-Aldrich) in PBS under constant agitation for 40 minutes at 37°C. This suspension was then centrifuged at 200xg for 5 minutes at room temperature. Next, the supernatant was aspirated and 10 ml of Trypsin-EDTA solution (T3924, Sigma-Aldrich) diluted 1/10 in PBS was added and incubated for 5 minutes at 37°C. Trypsin was then de-activated using 1 ml of Fetal Bovine Serum (FBS, FBS-12A, Capricorn Scientific GmbH) and the suspension was centrifuged at 200xg for 5 minutes at room temperature. The supernatant was then aspirated and the specimen was re-suspended in 10 ml Dulbecco's Modified Eagle Medium (DMEM, D6429, Sigma-Aldrich) followed by thorough trituration of the suspension using a Pasteur Pipette. The suspension was then sent through a cell strainer (40μm pore size, 352340, BD Falcon) and collected for centrifugation at 200xg for 5 minutes at room temperature. The cell pallet was re-

suspended in 10 ml DMEM and centrifugation was repeated. Medium was aspirated and cells were re-suspended in 1 ml of HBEC medium (components are listed in **Supplementary Table 1**) for counting and cell viability assessment using the Countess II Automated cell counter (Invitrogen). The cell yield of the PDX tumor models is shown in **Supplementary Table 2**.

Spheroid formation from isolated PDX tumor cells

For spheroid formation we adapted our published protocol[20] to make it suitable for PDX tumors. Briefly, we prepared agarose casts using the 12-256 precision micro-mold from Microtissues®. Isolated tumor cells were loaded onto the cast in a concentration between $2.0\text{-}4.0*10^6$/ml and were allowed to settle for 15 minutes. 2 ml of HBEC medium supplemented with 50 μg/ml of fibronectin (10838039001, Sigma-Aldrich) was then carefully added and spheroids were allowed to form for 5-7 days.

Spheroid treatments

After 5-7 days of spheroid formation, Olaparib (10 μM, S1060, Selleckchem) and radiation (10 Gy) treatments were performed and spheroids were harvested 7 days later. Cisplatin (10 μM) treatment was done 9-11 days post-formation and spheroids were harvested 3 days post-treatment. Each treatment was performed using a single agarose cast containing 256 spheroids.

Immunostaining

For staining of spheroids, 30 μM of EdU (ab146186, Abcam) was added to the spheroid samples for 2 hours at 37°C. Culture medium was then removed, and spheroids were harvested according to the manual from Microtissues®. After this, spheroids were transferred to a 1.5 ml low-binding microcentrifuge tube (MTL-0150-BC, Biotix) and centrifuged at 200xg for 5 minutes at room temperature. Supernatant was removed and 1 ml of PBS was added followed by centrifugation at 200xg for 5 minutes at room temperature. PBS was discarded and spheroids were fixed with 4% paraformaldehyde (PFA, 158127, Sigma-Aldrich) for 30 minutes at room temperature on a rotation wheel. Spheroids were then washed 3 times with PBS for 10 minutes

and centrifuged at 200xg for 5 minutes at room temperature after each wash step. After this, 1 ml of blocking solution (PBS supplemented with 1.0% (v/v) Triton X-100 (108603, Merck Millipore) and 3.0% (w/v) BSA (A9418, Sigma-Aldrich)) was added and incubated for 1 hour at room temperature on a rotation wheel. Spheroid samples were then centrifuged at 200xg for 5 minutes at room temperature after which blocking solution was removed and 1 mL of Click-IT reaction mix (components are listed in **Supplementary Table 3**) was added. This was then incubated for 2 hours at room temperature in the dark on a rotation wheel. After incubation, samples were centrifuged and washed twice with 1 ml of blocking solution for 10 minutes and centrifuged at 200xg for 5 minutes at room temperature after each wash step. After this 1 ml of blocking solution was added to the spheroids and samples were incubated for 4 hours at room temperature on a rotation wheel before incubation with the primary antibodies (the sources and dilutions are listed in **Supplementary Table 4**) diluted in blocking solution for a duration of 48 hours at 4°C on a rotation wheel. After this, spheroid samples were washed 3 times for 10 minutes with PBS supplemented with 0.1% (w/v) BSA at room temperature on a rotation wheel before a 3-hour incubation with the secondary antibodies (the sources and dilutions are listed in **Supplementary Table 5**) diluted in blocking solution at 4°C. After this, samples were washed 2 times for 15 minutes with PBS at room temperature before mounting the spheroids using Vectashield containing 4',6-diamidino-2-phenylindole (DAPI, H-1200, VectorLabs) for staining of the nucleus.

Bright-field microscopy

To monitor spheroid sizes, bright-field images were captured using the Olympus IX70 inverted phase contrast microscope. Images were captured using a DFC320 digital camera from Leica Microsystems and saved using Adobe Photoshop 2021. We used the UPlanFLN 4x PhL objective (pixel size, 0.183 mm) for all CS spheroid models and the UPlanFL 10x objective (pixel size, 0.123 mm) for the CR models.

Confocal microscopy

Stained spheroids were imaged using the SP8 confocal microscope from Leica Microsystems with LasX suite control software. We used Leica's air immersion HC PL APO 20X/0.75 CS2 objective on the SP8. Sequential scanning was employed to prevent crosstalk when detecting multiple fluorophores in a single sample. The following excitation and filter settings were applied: 405 nm excitation, 415-470 nm for detection of DAPI nuclear stain; 488 nm excitation, 500-550 nm for detection of Alexa Fluor 488 and 561 nm excitation, 570-620 nm for detection of Alexa Fluor 555.

Image analysis

EdU+ cells and 53BP1 foci in spheroids and geminin+ cells and RAD51 foci in tissue slices were quantified using the Image-J macro "Foci-analyzer" (freely available at https://github.com/BioImaging-NKI/Foci-analyzer; created by Bram van den Broek, the Netherlands Cancer Institute, the Netherlands)[31]. Cleaved caspase-3 signal in spheroids was quantified using ImageJ by manually selecting the ROI per Z-slice based on the DAPI signal. The mean intensity of cleaved caspase-3 signal was then measured in this region.

Model Training

We implemented the Segment Anything Model (SAM) for microcopy framework for our automated spheroid segmentation. Training data was generated using the micro-SAM napari plugin (version 1.0.1), with spheroids manually annotated using rectangular prompts. For preparation of training, images were converted to 8-bit grayscale format in Fiji. The training dataset comprised 96 images (2 images of each model, time point and treatment) with corresponding manual annotations, split equally between training and validation sets. The semi-manual segmentation predictions based on the prompt were saved as label images.

For fine-tuning the Segment Anything Model for automatic instance segmentation we used a Segment Anything for Microscopy notebook (https://github.com/computational-cell-

analytics/micro-sam/notebooks/sam_finetuning.ipynb) as reference and modified this notebook to work with our data set and to ease training of different models using subsets of data. We also added a script to process the data and to extract several metrics from the SAM generated label images. The model was trained for 100 epochs, 5 objects per batch, a patch size of 1024x1024 pixels and using the additional convolutional decoder for end-to-end automatic instance segmentation[32].

Part of the data of the resistant tumors was acquired on a different microscope. To allow for training the canvas of all images were set to the dimensions of the largest image in the data set, resulting in a zero-value border around the images and label images.

Training was performed on an Ubuntu 20.04 system with an Intel Xeon Gold4218 (64 cores) processor, 512 GB memory, and a Quadro RTX 6000 24GB GPU.

Post-segmentation analysis included automated measurement of multiple spheroid parameters including their individual area, height, width, perimeter, diameter, and volume.

Evaluation of network performance

Network performance was evaluated by precision, recall and F1 score. True positives (TP) were defined as spheroids correctly measured; true negatives (TN) were defined as empty wells. False positives (FP) were defined as incorrectly measured spheroids which showed a too large or too small surface area based on overlay of AI-recognized and bright-field image. Precision, Recall and F1 scores were calculated based on these parameters.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \qquad (1)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \qquad (2)$$

$$\text{F1 score} = \frac{2\,\text{Recall} \times \text{Precision}}{\text{Recall+Precision}} \qquad (3)$$

Inclusion and ethics

All experiments were approved by the Central Committee on Animal Experiments (CCE) and performed in accordance with the guidelines of the Institute for Animal Welfare of the Netherlands Cancer Institute. We declare no inclusion and ethics concerns related to this study.

Statistics and Reproducibility

Statistical analysis and data visualization was performed using GraphPad Prism v10.2.1 (GraphPad Software Inc.). Error bars represent standard deviation (SD) over data points. Statistical analysis was conducted using at least three independent experiments with Student's T-tests (unpaired, parametric two-tailed tests with a 95% confidence interval), with or without Welch's correction, or one/two-way Anova tests. A $P<0.05$ was considered statistically significant. Four tumor models (biologically replicates) from three different mice were used to form spheroids (CS1-3, CR1-3 with CS2A and CS2B and CR2A and CR2B extracted from mice with two tumors). The degree of correlation was determined by calculating Pearson's correlation coefficients. To set a threshold for our prediction analysis, outliers were identified in the control condition using the Robust regression and Outlier removal (ROUT) method in GraphPad Prism. For the ROUT analysis we compared results for various Q values: Q=0.1%, Q=1%, Q=5% and Q=10%. The threshold was determined based on the mean fold change spheroid area (untreated, day 7; all CS models combined) and the standard deviation (SD).

$$Threshold = mean(untreated) - 2 \times SD(untreated)$$

Results

Treatment response of PDX spheroids faithfully reflects expected *in vivo* response

To develop a spheroid assay to report on treatment response in breast cancer, we used the cisplatin-sensitive PDX model T250 (TNBC), which is HR deficient due to *BRCA1* mutations (c.2210delC/ 2329delC)[30]. Cisplatin-resistant (CR) tumors were established from the cisplatin-sensitive T250 (CS) tumors by repeated *in vivo* cisplatin treatment as described previously[30].

The HR status of CS tumors was confirmed using the RECAP assay[14,17]. In contrast, CR tumors displayed on average in 75% of nuclei of geminin+ cells >5 RAD51 foci, indicating that these tumors were HRP explaining their cisplatin resistance (**Supplementary Fig. 1**). Tumor tissue slices were stained with 53BP1 to confirm the induction of DNA double strand breaks (DSBs) after irradiation (**Supplementary Fig. 1a, b**).

We used an agarose mold system to produce spheroids of CS and CR tumors, resulting in 256 spheroids for each condition we tested. CS and CR spheroids were formed from tumors from three different mice (CS1-3; CR1-3, respectively) with mice carrying two tumors (CS2A and CS2B; CR2A and CR2B, respectively). The formed spheroids displayed a consistent morphology with a size of 150-300 µm in diameter (**Fig. 1a**) and consisted mainly of tumor cells (pan-cytokeratin+ cells; **Supplementary Fig. 2a, b**). Additionally, we confirmed CS and CR spheroids contained proliferative cells (EdU+, a S-phase marker) throughout the whole spheroid, suggesting that spheroids did not contain a necrotic core (**Supplementary Fig. 2c**). We tested the response to cisplatin, the PARP inhibitor olaparib and irradiation in these CS and CR spheroids. **Fig. 1a** and **b** show a schematic representation of spheroid formation and our experimental approach, respectively.

Treatment of CS tumors with cisplatin, olaparib and radiation induced apoptosis as observed by significantly increased cleaved caspase-3 staining intensity compared to untreated spheroids (**Fig. 1e**). Apoptosis was present throughout the spheroid, suggesting cisplatin and olaparib penetrated multiple cell layers and reached cells at the core (**Supplementary Fig. 2d**). Cell proliferation was significantly reduced upon treatment with olaparib or radiotherapy, but more EdU+ cells were present in CS spheroids after cisplatin treatment (**Fig. 1c, f**), most likely because cisplatin may slow down S-phase progression [33,34]. CR spheroids did not respond to cisplatin and olaparib as observed by the lack of apoptosis (**Fig. 1d, g**). These results were expected, as *in vivo* resistance to cisplatin occurs in CR models[30] through restoration of HRR

(**Supplementary Fig. 1**), which should result in reduced PARPi sensitivity. Furthermore, olaparib-treated CR spheroids showed no significant difference in the number of EdU+ cells compared to the control (**Fig. 1h**). In contrast, cisplatin-treated CR spheroids showed a significant increase in EdU+ cells similar to cisplatin-treated CS spheroids. We observed a similar response to irradiation in CR spheroids compared to CS spheroids (**Fig. 1e-h**), which was expected as most DSBs induced by irradiation are repaired via non-homologous end-joining[35] and therefore HRD status of tumors would have less of an effect on treatment response. In CS spheroids, we found that all treatments resulted in significantly more DSBs (indicated by 53BP1 foci) compared to the untreated condition, although the percentage of nuclei with >5 foci varied between spheroids in the same condition (**Fig. 1i-j**).

Overall, CS and CR spheroids faithfully displayed the expected *in vivo* responses to the respective treatments. However, we noticed that both the degree of apoptosis and inhibition of cell proliferation induced by the treatments varied substantially between individual spheroids. For instance, olaparib-treated CS and CR spheroids displayed a lower or higher degree of apoptosis or proliferation in the same biological sample (**Supplementary Fig. 3a, b**) and the fold mean intensity of cleaved caspase-3 and number of EdU+ cells varied substantially between spheroids in treated conditions of CS and CR tumor models (**Supplementary Fig. 3c, d** and **e, f,** respectively). We therefore aimed to develop an alternative assay that would allow us to capture the heterogeneity of the spheroids, by monitoring a large number of spheroids at multiple time points, using only a bright-field microscope.

Automated spheroid segmentation and size determination using an AI model

Spheroid size is influenced both by cell death and cell proliferation and therefore may be a suitable readout for treatment response. An additional advantage of monitoring spheroid size is that it requires less sample preparation compared to staining and allows longitudinal measurements of the same spheroids at different time points.

High-throughput analysis of spheroid size requires automated segmentation. For automatic segmentation of the spheroids from the bright-field images we explored different options that would segment spheroids from the images in an unbiased way. Traditional intensity segmentation methods are difficult on bright-field images given that the wells, in which the spheroids grow, are clearly visible. Subsequently we have tested a random-forest segmentation approach, for which a few images are sparsely annotated to create a pixel classification model. While this approach worked relatively well in some of our datasets, it was difficult to create a model that would work on all images accurately. We therefore decided to use a python implementation of SAM dedicated for light and electron microscopy named Segment Anything for Microscopy[32] which allows model fine-tuning for automatic instance segmentation using an additional image encoder. Models that were not fine-tuned on our data were not able to accurately segment the spheroids. However fine-tuning of the model on our images resulted in much better segmentation of the spheroids (**Supplementary Fig. 4**). Specifically, we generated a training set for micro-SAM using semi-manual annotation, by only marking rectangles around the spheroids in the images and used SAM to segment the spheroids. We then trained SAM on those annotated images to automatically segment the spheroids from all images without the need of manual annotation. **Fig. 2a** shows the pipeline for automatic detection of spheroids in bright-field images.

For validation and testing of the robustness of the segmentation, we fine-tuned the SAM model using different subsets of our data (**Fig. 2a**). We used image label sets of individual CS tumors and a combined dataset from all CS tumors to train the model and compared the results. In all cases, we used two images, randomly selected from each treatment condition and time point of the respective CS tumor to have an unbiased representation of spheroids. We used one set of images as training data and the other set for validation during training. We manually inspected all segmentations and assessed the accuracy of the models by calculating F1 scores

(determined by the model's recall and precision) for each training approach for the four CS tumors. To calculate the recall, we defined true positives as correctly-segmented spheroids and true negatives as empty wells. We hardly found true negatives, resulting in a recall close to 1.0. We defined false positives as incorrectly segmented spheroids based on the overlay of spheroids recognized by AI and the bright-field image (examples of overlays in **Supplementary Fig. 5a-d**). We found high F1 scores for CS models trained on its own images (>0.95); however, when trained on one tumor model and applied to a different model, lower F1 scores were observed for some models (**Fig. 2b**). When trained on all CS models combined, the obtained F1 scores were as high as when trained on its own images (**Fig. 2b-d**). The F1 scores were very similar for the different time-points, suggesting that treatment did not influence segmentation results (**Fig. 2d**).

Additionally, we manually measured the horizontal and vertical cross-sectional distances of spheroids for CS1, CS2A and CS2B tumors and calculated the spheroid diameter per image by averaging these two measurements. We then compared the manually- and AI-determined diameters and calculated the degree of correlation (**Fig. 2e**, **Supplementary Fig. 5h-i**). We found on average a Pearson's r=0.7457+/-0.1025 when the model was trained using its own images, and lower correlations when the model was trained with images from a different CS tumor (average r=0.5512+/-0.081; **Fig. 2g**). Also, the fold-change in the AI-measured area after treatment correlated well with the manually-measured area (**Fig. 2f**). Again, when we used random images from the four CS tumors together, we observed a similar correlation as when trained on images of the same tumor alone (average r=0.7249+/-0.077). Considering the pixel size in the bright-field images is circa 1.5 µm, the manually determined height and width of spheroids cannot be considered an absolute ground truth simply because manually measuring spheroid dimensions is not that accurate. In conclusion, our data suggests that training of the AI model using images from all four CS tumors was as accurate as using its own images for

training. We therefore decided to use the AI model trained with images from all four CS tumors together to analyze the spheroid measurements.

Tumor volumes are commonly used to determine treatment response[26]. However, our bright-field images do not provide a 3D image of spheroids, so we need to assume that spheroids are spheres. To validate that spheroids in 2D images are indeed round, we determined the degree of roundness based on 1) the length to width ratio of the box that can be placed directly around the spheroid area (output of the AI model) and 2) the linear correlation between measured (total number of pixels) and circular surface area (based on the diameter). We found that for most spheroids, the normalized height to width ratio approaches 1.0, indicating that most spheroids were indeed round (**Fig. 2h**). Moreover, normalized calculated and measured surface areas correlated well for the four CS tumors, independent of spheroid treatment (**Fig. 2i-j, Supplementary Fig. 5e-g**). We therefore decided to use the measured surface areas for further analysis.

We next tested if our AI model trained on all four CS tumors could accurately determine the treatment response of CR tumors, a model for *in vivo* acquired resistance to cisplatin. We compared the spheroid measurements obtained from automatic spheroid segmentation using different training approaches, namely trained on the four CS models together (CSall), the four CR models together (CRall) or CSall combined with CRall (CSCR). Treatment response of CR2B and CR3 spheroids showed very similar trends when comparing the different training approaches (**Supplementary Fig. 6a-c** and **6d-f**, respectively). The calculated F1 scores suggest that the AI model trained on CSall performed the best (**Supplementary Fig. 6g-i**). We therefore decided to use this AI model for the analysis of treatment response.

Spheroid measurements report on the expected treatment response of CS and CR tumors

To validate whether our assay could accurately report on treatment response, we analyzed

bright-field images from untreated and CS and CR spheroids treated with olaparib, cisplatin and radiation and measured their surface area. **Fig. 3a** and **b** show an example of bright-field images of untreated and treated CS and CR spheroids at day 0, 4 and 7 in culture, respectively. Measured surface areas for each condition were normalized to their corresponding average area on day 0. In all four CS tumors we found that olaparib-treated spheroids displayed surface areas that mostly fell below the 25th percentile of the area distribution of untreated spheroids (violin plots; **Fig. 3c-f**), suggesting most spheroids responded to the treatment (also indicated by a significant decrease in surface area in olaparib-treated compared to untreated spheroids; **Supplementary Fig. 7a-d**). In contrast, olaparib-treated CR spheroids were similar in size or larger than untreated spheroids, indicating that they did not respond to the treatment (**Fig. 3g-h, j; Supplementary Fig. 7e-f**, **h**), except CR2B spheroids (**Fig. 3i; Supplementary Fig. 7g**, *P* =0.0106). In all four CS tumors, olaparib-treated spheroids displayed the most reduced surface area after treatment, which was expected considering the degree of cell death and effect on cell proliferation that we found earlier (**Fig. 1c, e, f**). However, the olaparib response is variable: CS1 responded well (**Fig. 3c**) whereas for CS2A a fraction of spheroids fell in the range of the untreated condition (**Fig. 3d**). Overall, CS spheroids displayed sensitivity to olaparib whereas CR spheroids appeared to be resistant, which may be explained by restoration of HRR capacity in these tumors (**Supplementary Fig. 1b**).

The response of CS and CR spheroids to radiotherapy was mixed. Most CS tumors responded to radiotherapy (**Fig. 3c-d**, **f**; **Supplementary Fig. 3a, c-d**), except for CS2A (**Fig. 3e**; **Supplementary Fig. 7b**). Overall, CR1 and CR3 spheroids seemed to be resistant (**Fig. 3g**, **j**; **Supplementary Fig. 7e**, **h**) as these spheroids were similar in size or larger compared to untreated spheroids, although a fraction of CR1 spheroids did respond, whereas CR2A and CR2B spheroids were significantly smaller than untreated spheroids (**Supplementary Fig. 7f**, **g**). Overall, both CS and CR spheroids showed a partial response to radiotherapy.

The effect of cisplatin treatment was the most variable. CS1 and CS3 spheroids significantly reduced in size post treatment, but cisplatin-treated C2A and CS2B spheroids were similar in size as untreated spheroids (**Supplementary Fig. 7a-d**). For CS1, cisplatin treatment revealed both a responding and non-responding population (**Fig. 3c**), whereas most CS2A and CS2B spheroids did not respond to cisplatin (**Fig. 3d** and **e**). A possible explanation for this observation may be that CS1 spheroids contained more proliferating cells at the start of the cisplatin treatment, as indicated by the increase in spheroid size between day 0 and day 4 (**Supplementary Fig. 7a**). This contrasted with CS2A and CS2B, which showed less size increase, suggesting less cell proliferation (**Supplementary Fig. 7b, c**). Most CR spheroids did not respond to cisplatin treatment for CR1, CR2A and CR3 tumors (**Supplementary Fig. 7e-f, h**). Only CR2B showed a significant decrease in spheroid size in cisplatin-treated versus untreated spheroids (**Supplementary Fig. 7g**, P=0.0450).

When comparing cisplatin-treated to olaparib-treated CR spheroids, we found either significantly larger olaparib-treated (**Fig. 3g**) or significantly larger cisplatin-treated spheroids (**Fig. 3h, j**). These findings reflect the results of our initial experiments, in which we did not observe an increase in cell death and reduction in cell proliferation compared to untreated spheroids when staining for cleaved caspase 3 and EdU, respectively (**Fig. 1g, h**). Overall, we observed differences in response to cisplatin between CS and CR spheroids. However, the cisplatin-response in CS spheroids seems to be only partial. A possible explanation for this is that the chosen cisplatin concentration may be too low for the majority of CS spheroids to reduce in size. It seems that the cisplatin concentration is high enough to induce apoptosis in CS spheroids, but does not seem to have a large enough effect to reduce spheroid size as non-responsive tumor cells may remain proliferative.

In conclusion, the CR spheroids responded differently to olaparib and cisplatin compared to CS spheroids, suggesting that measurements of spheroid surface area can report on the tumor's

expected *in vivo* response. Moreover, our results suggest that CS tumors on average responded the most to olaparib treatment.

Discriminating between treatment-sensitive and resistant tumors

We have shown that our assay is able to report on differences in treatment response between CS and CR spheroid models. As an example, **Fig. 4a** shows a clear difference in changes in surface area between CS1 and CR1 spheroids upon treatment with olaparib. Next, we wanted to find a threshold of spheroid area fold-change to discriminate between sensitive and resistant tumors. We first analyzed the normalized frequency distribution of spheroid areas in untreated CS models (all data combined). This frequency distribution follows a Gaussian distribution with the top of the curve close to 1.0 (**Supplementary Fig. 8a**). The areas of untreated CR spheroids followed a similar Gaussian distribution (**Supplementary Fig. 8b**), suggesting that the datasets are very similar for the two tumor models. The frequency distribution of olaparib-treated CS1 spheroids shifted to the left compared to CR1, suggesting that CS1 spheroids responded to the treatment (**Fig. 4b**). Cisplatin- and radiation-treated spheroids displayed frequency distributions that overlap, indicating a partial response (**Fig. 4c, f**) as we have previously seen (**Fig. 3**). To determine the threshold value for sensitivity, we first identified and removed outliers using the Robust regression and Outlier removal (ROUT) method[36] and then determined the distribution of spheroids sizes to be expected for the untreated condition (example of Q=10% in **Supplementary Fig. 8c**). We compared four different Q values for outlier detection: Q=0.1% (mild), Q=1% (commonly used), Q=5% and Q=10% (more aggressive). Next, we calculated the mean value and standard deviation for the cleaned-up datasets, filtered for outliers, to set the threshold 2SD lower than the mean (to have a confidence interval of 95%). We used this threshold to define responders (below threshold) and non-responders (above threshold) for treated CS and CR models (shown for Q=10%; **Fig. 4d, e** and **Supplementary Fig. 8d, e**). We used the threshold value from CS models to apply to the CR data. The

percentage of spheroid data points measured below the calculated threshold for CS and CR tumor models is shown in heatmaps (Fig. 4g and h, respectively for Q=10%; Supplementary Fig. 9a-f) for Q=0.1%, Q=1% and Q=5%), indicating the prediction of response. For Q=0.1%, Q=1% and Q=5% we observed the same trends in prediction of response for CS and CR models. However, with Q=10% we found a better distinction for olaparib sensitivity between CS and CR models compared to Q=1% (**Supplementary Fig. 9g and h**). **Fig. 4g** and **h** show the prediction of responders for CS and CR tumors, respectively (with Q=10%). We observed a clear difference in the prediction of sensitivity for olaparib between the CS and CR models (on average 78.3% and 13.9% of responding spheroids, respectively). We did not observe a significant difference in percentage of responding spheroids between the CS and CR models after radiotherapy (**Supplementary Fig. 9e,** Q=10%), which was to be expected as we did not find a clear difference in response based on our staining results (**Fig. 1**) and spheroid measurements (**Fig. 3**). We also did not find a significant difference in percentage of responders between the CS and CR models after cisplatin treatment (**Supplementary Fig. 9e,** Q=10%). This observation may be explained by the partial response to cisplatin we observed in the CS spheroids. Moreover, these results suggest that although the average response may show sensitivity, a substantial fraction of spheroids may not respond, emphasizing the variation in tumor response. Further, based on the comparison of olaparib, cisplatin and radiation treatment, we can conclude that olaparib would be the best treatment choice for these CS tumors.

Discussion

In this study we developed a high-throughput assay for functional breast cancer precision medicine based on spheroid sizes as a measure for treatment response. Functional precision oncology may be extremely valuable for cancer treatment outcome, but major challenges need to be overcome to implement assays in the clinic. Technological challenges include the

limitations of models to mimic the complex tumor micro-environment (TME), the limited amount of patient sample for testing, and the limited readouts for treatment prediction[37].

Spheroids, organoids and tissue explants have been considered as models for functional precision oncology, with each their own advantages and limitations[12]. Spheroid and organoid formation requires cell isolation from the tumor, destroying tissue architecture and the native TME in contrast to tumor explants[12]. Unlike organoids, spheroids do not spontaneously recapitulate the native TME[12]. Spheroid models may include additional compartments of the TME, for instance by forming mixed spheroids from isolated tumor, stromal and immune cells[38]. Alternatively, organotypic spheroids may be used, consisting of partially dissociated tumors and preserving more of the native TME[39]. In our study we used spheroids containing only tumor cells, which was sufficient to correlate the treatment response of CS and CR spheroids with the expected *in vivo* treatment response in mice. However, our assay is compatible with these more complex spheroid models.

In most cases, biopsies of patient tumors will be the only material available for functional assays and hence the amount of material is limited, posing challenges for spheroid-based functional assays. As spheroids are formed and treated within a brief duration, expansion of the material is not possible. As we formed and treated 256 spheroids per condition, but only extracted data from approximately 60 spheroids for each tumor sample, theoretically we could use 75% less starting material when we would be able to analyze all formed spheroids. In the future the assay used in this study needs to be optimized to show whether it is compatible with restricted material from tumor biopsies from patients. Alternatively, patient-derived organoids (PDOs) are suitable for long-term culturing and passaging, allowing to create living biobanks of patient tumor models[40] and have been used to perform drug screens and for immune-oncology[41]. Breast cancer PDOs recapitulate disease heterogeneity, typically match histopathology, hormone receptor status and HER2 receptor status and show similar responses *in vitro* compared to their

orthotopic transplant in mice[40]. However, for clinical practice, results are required within a brief time span (2 weeks) to not delay treatment for the patient. This poses challenges to use organoids for this purpose as the generation of PDO's is time-consuming (multiple weeks or months)[42]. In contrast, spheroid formation is fast, and we were able to form spheroids and start treatment 5-7 days after cell isolation.

For functional precision oncology to be successfully applied to the clinic, functional assays should provide a clear readout and preferably be reasonably effortless and low cost. Moreover, due to tumor heterogeneity and variation in readouts, functional assays may become more reliable with larger numbers of data-points. Spheroid cultures are well-suited for high-throughput assays, especially with the recent development of protocols for large scale spheroid production[20,21,22]. The challenge of high-throughput functional assays lies in developing suitable readouts. Tumor volume assessment is one of the most used parameters to determine treatment efficacy[26]. In this study, we used spheroid size as readout for treatment response which is influenced by cell death and proliferation and allows for longitudinal monitoring of the same spheroids at different time points. By using deep learning to implement automatic spheroid segmentation, the analysis pipeline we established is suitable for large-scale drug screening. Overall, spheroid measurements reported accurately on sensitivity to treatment. Especially for the olaparib treatment, a clear distinction could be made between sensitive and resistant tumor models. However, in this study CS spheroids showed a partial response to cisplatin based on spheroid size (**Fig. 3**), whereas the staining results for apoptosis displayed a significant difference between cisplatin-treated CS and CR spheroids (**Fig. 1**). In contrast, the number of EdU+ cells was significantly increased in both cisplatin-treated CS and CR spheroids (**Fig. 1**). These observations may be explained by treatment with an insufficient dose of cisplatin resulting in a cell killing effect while a fraction of less-responsive tumor cells remains proliferating. Whole-mount staining of spheroids for apoptosis markers (**Fig. 1**) is a labor-

intensive assay, requiring confocal imaging and analysis. To perform this assay on large numbers of spheroids is not easily implemented in the clinic. However, addition of an apoptosis readout, such as a live/dead staining, to spheroid measurements would provide more insight into the treatment response and would at least partially overcome the limitations posed by only using bright-field images of spheroids as readout for treatment response. Our assay may be compatible with incorporating such a fluorescent probe after optimization. This approach would require the use of high-content screening microscopy combining fluorescence with bright-field imaging. Incorporation of fluorescent intensity measurements should be compatible with the automatic spheroid segmentation after adaptation of the imaging analysis program. Furthermore, future studies could leverage changes in the morphology of the spheroids, extracted from the bright-field images. Combined with the size measurements we have already implemented, this might give additional insights in the response of the spheroids to different treatments. Alternatively, if spheroids would be cultured in ultralow attachment plates, as a single spheroid per well, than ATP-based cell viability assays such as Cell Titer Glo 3D, could be used in addition to the automated spheroid segmentation to obtain more insight into the treatment response.

To facilitate large-scale assessment of spheroid measurements, automatic segmentation of spheroids is valuable, although challenging. The nature of images obtained from our spheroid cultures complicates automatic segmentation due to variability in signal intensities and shapes of spheroids. Although in future studies image variability could be reduced by automatic image acquisition, the assay requires robust tools to quantify the images. While classical segmentation methods based on signal thresholds and circularity[27], or pixel classifiers such as ilastik did not provide reliable segmentation in our hands, deep learning-based approaches offered a solution. By implementing a microscopy-optimized version of the Segment Anything model, we achieved robust automated spheroid segmentation across different tumor models. The high F1 scores

(>0.9) obtained for both training (CS) and test (CR) models demonstrate the broad applicability of this approach without the need for model-specific training. Our automated analysis pipeline enables efficient processing of large-scale spheroid screening data, offering many potential applications. In the future, the assay may be further automated by implementing automated microscopy to increase sensitivity and make assays less labor-intensive.

Conclusion

In summary, we describe a high-throughput treatment prediction assay using spheroids and automated analysis of treatment responses using deep learning, which can be adapted for biopsies derived from patient tumors for functional precision oncology. We used PDX models to develop and validate our spheroid assay, but these models may introduce a biological mismatch between host organism and xenograft, limiting their biological representation of a patient's tumor[12,43]. Therefore, further investigation of the potential value of this assay for functional precision oncology requires studying the correlation between treatment response of spheroids formed from resected/biopsy material and patient outcome.

Data Availability

Training data and fine-tuned models are accessible through Zenodo (https://doi.org/10.5281/zenodo.14832406)[44]. The numerical data plotted (source data) in Figs. 1, 2, 3, and 4 are provided in Supplementary Data 1 and 2.

Code Availability

All code, including the training notebook and analysis scripts, is available on GitHub (https://github.com/maartenpaul/spheroidAnalysis; https://doi.org/10.5281/zenodo.15239560)[45]. Code development was assisted by the Anthropic Claude 3.5 Sonnet large language model. All code was thoroughly validated and verified by the authors. We take full responsibility for its functionality and results.

Author contributions

Conceptualization: M.M.P.K. Funding acquisition: R.K and D.vG. Investigation: B.H, J.J.T, M.B and M.M.P.K. Methodology: B.H, M.W.P, M.M.P.K. Visualization: M.W.P, B.H and M.M.P.K. Supervision: R.K, J.J, D.vG and M.M.P.K. Data curation: M.B and Z.M.K. Software: M.W.P. Formal analysis: B.H, M.W.P and M.M.P.K. Writing: M.W.P and M.M.P.K. Review and editing: M.W.P, R.K, J.J, D.vG and M.M.P.K.

Declaration of interests

The authors declare no competing interests.

References

1. Arnold, M. *et al.* Current and future burden of breast cancer: Global statistics for 2020 and 2040. *Breast* **66**, 15–23 (2022).

2. Newman, L. Oncologic anthropology: Global variations in breast cancer risk, biology, and outcome. *J. Surg. Oncol.* **128**, 959–966 (2023).

3. Lund, M. J. *et al.* Race and triple negative threats to breast cancer survival: A population-based study in Atlanta, GA. *Breast Cancer Res. Treat.* **113**, 357–370 (2009).

4. Hartman, A.-R. *et al.* Prevalence of BRCA mutations in an unselected population of triple-negative breast cancer. *Cancer* **118**, 2787–95 (2012).

5.   Heeke, A. L. *et al.* Prevalence of Homologous Recombination–Related Gene Mutations Across Multiple Cancer Types. *JCO Precis. Oncol.* 1–13 (2018) doi:10.1200/po.17.00286.

6.   van Wijk, L. M., Nilas, A. B., Vrieling, H. & Vreeswijk, M. P. G. RAD51 as a functional biomarker for homologous recombination deficiency in cancer: a promising addition to the HRD toolbox? *Expert Rev. Mol. Diagn.* **22**, 185–199 (2022).

7.   Melinda, L. T. *et al.* Homologous recombination deficiency (hrd) score predicts response to platinum-containing neoadjuvant chemotherapy in patients with triple-negative breast cancer. *Clin. Cancer Res.* **22**, 3764–3773 (2016).

8.   Mangogna, A. *et al.* Homologous Recombination Deficiency in Ovarian Cancer: from the Biological Rationale to Current Diagnostic Approaches. *J. Pers. Med.* **13**, (2023).

9.   Joosse, S. A. *et al.* Prediction of BRCA2-association in hereditary breast carcinomas using array-CGH. *Breast Cancer Res. Treat.* **132**, 379–389 (2012).

10. Davies, H. *et al.* HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational signatures. *Nat. Med.* **23**, 517–525 (2017).

11. Nguyen, L., W. M. Martens, J., Van Hoeck, A. & Cuppen, E. Pan-cancer landscape of homologous recombination deficiency. *Nat. Commun.* **11**, 1–12 (2020).

12. van Renterghem, A. W. J., van de Haar, J. & Voest, E. E. Functional precision oncology using patient-derived assays: bridging genotype and phenotype. *Nat. Rev. Clin. Oncol.* **20**, 305–317 (2023).

13. Stover, E. H., Fuh, K., Konstantinopoulos, P. A., Matulonis, U. A. & Liu, J. F. Clinical assays for assessment of homologous recombination DNA repair deficiency. *Gynecol. Oncol.* **159**, 887–898 (2020).

14. Meijer, T. G. *et al.* Functional RECAP (REpair CAPacity) assay identifies homologous recombination deficiency undetected by DNA-based BRCAness tests. *Oncogene* **41**, 3498–3506 (2022).

15. Naipal, K. A. T. *et al.* Functional Ex vivo assay to select homologous recombination-deficient breast tumors for PARP inhibitor treatment. *Clin. Cancer Res.* **20**, 4816–4826 (2014).

16. van Wijk, L. M. *et al.* The recap test rapidly and reliably identifies homologous recombination-deficient ovarian carcinomas. *Cancers (Basel).* **12**, 1–16 (2020).

17. Meijer, T. G. *et al.* Functional ex vivo assay reveals homologous recombination deficiency in breast cancer beyond BRCA gene defects. *Clin. Cancer Res.* **24**, 6277–6287 (2018).

18. Kapałczyńska, M. *et al.* 2D and 3D cell cultures – a comparison of different types of cancer cell cultures. *Arch. Med. Sci.* **14**, 910–919 (2018).

19. Costa, E. C., Gaspar, V. M., Coutinho, P. & Correia, I. J. Optimization of liquid overlay technique to formulate heterogenic 3D co-cultures models. *Biotechnol. Bioeng.* **111**, 1672–1685 (2014).

20. Haspels, B. & Kuijten, M. M. P. Protocol for formation, staining, and imaging of 3D breast cancer models using MicroTissues mold systems. *STAR Protoc.* **5**, 103250 (2024).

21. Grexa, I. *et al.* SpheroidPicker for automated 3D cell culture manipulation using deep learning. *Sci. Rep.* **11**, 1–11 (2021).

22. Trossbach, M., Åkerlund, E., Langer, K., Seashore-Ludlow, B. & Joensson, H. N. High-throughput cell spheroid production and assembly analysis by microfluidics and deep learning. *SLAS Technol.* **28**, 423–432 (2023).

23. Mukundan, S. *et al.* Automated Assessment of Cancer Drug Efficacy On Breast Tumor Spheroids in Aggrewell™400 Plates Using Image Cytometry. *J. Fluoresc.* **32**, 521–531 (2022).

24. Linxweiler, J. *et al.* Patient-derived, three-dimensional spheroid cultures provide a versatile translational model for the study of organ-confined prostate cancer. *J. Cancer Res. Clin. Oncol.* **145**, 551–559 (2019).

25. Della Corte, C. M. *et al.* Antitumor activity of dual blockade of PD-L1 and MEK in NSCLC patients derived three-dimensional spheroid cultures. *J. Exp. Clin. Cancer Res.* **38**, 1–12 (2019).

26. Piccinini, F., Tesei, A., Arienti, C. & Bevilacqua, A. Cancer multicellular spheroids: Volume assessment from a single 2D projection. *Comput. Methods Programs Biomed.* **118**, 95–106 (2015).

27. Celli, J. P. *et al.* An imaging-based platform for high-content, quantitative evaluation of therapeutic response in 3D tumour models. *Sci. Rep.* **4**, 1–10 (2014).

28. Kirillov, A. *et al.* Segment Anything. *Proc. IEEE Int. Conf. Comput. Vis.* 3992–4003 (2023) doi:10.1109/ICCV51070.2023.00371.

29. Benning, L., Peintner, A., Finkenzeller, G. & Peintner, L. Automated spheroid generation, drug application and efficacy screening using a deep learning classification: a feasibility study. *Sci. Rep.* **10**, 1–11 (2020).

30. Ter Brugge, P. *et al.* Mechanisms of therapy resistance in patient-derived xenograft models of brca1-deficient breast cancer. *J. Natl. Cancer Inst.* **108**, 1–12 (2016).

31. Koob, L. *et al.* MND1 enables homologous recombination in somatic cells primarily outside the context of replication. *Mol. Oncol.* **17**, 1192–1211 (2023).

32. Archit, A. *et al. Segment Anything for Microscopy*. *Nature Methods* vol. 22 (Springer US, 2025).

33. Wang, P. *et al.* Cisplatin induces hepG2 cell cycle arrest through targeting specific long noncoding RNAs and the p53 signaling pathway. *Oncol. Lett.* **12**, 4605–4612 (2016).

34. Wagner, J. M. & Karnitz, L. M. Cisplatin-induced DNA damage activates replication checkpoint signaling components that differentially affect tumor cell survival. *Mol. Pharmacol.* **76**, 208–214 (2009).

35. Mahaney, B. L., Meek, K. & Lees-Miller, S. P. Repair of ionizing radiation-induced DNA double strand breaks by non-homologous end-joining. *Biochem J.* **417**, 639–650 (2009).

36. Motulsky, H. J. & Brown, R. E. Detecting outliers when fitting data with nonlinear regression - A new method based on robust nonlinear regression and the false discovery rate. *BMC Bioinformatics* **7**, 1–20 (2006).

37. Ayuso, J. M., Virumbrales-Muñoz, M., Lang, J. M. & Beebe, D. J. A role for microfluidic systems in precision medicine. *Nat. Commun.* **13**, 1–12 (2022).

38. Yin, S. *et al.* Patient-derived tumor-like cell clusters for drug testing in cancer therapy.

Sci. Transl. Med. **12**, (2020).

39.   Aref, A. R. *et al.* 3D Microfluidic Ex Vivo Culture of Organotypic Tumor Spheroids to Model Immune Checkpoint Blockade. *Lab Chip* **18**, 3129–3143 (2019).

40.   Sachs, N. *et al.* A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity. *Cell* **172**, 373-386.e10 (2018).

41.   Kretzschmar, K. Cancer research using organoid technology. *J. Mol. Med.* **99**, 501–515 (2021).

42.   Dekkers, J. F. *et al.* Long-term culture, genetic manipulation and xenotransplantation of human normal and breast cancer organoids. *Nat. Protoc.* **16**, 1936–1965 (2021).

43.   Ireson, C. R., Alavijeh, M. S., Palmer, A. M., Fowler, E. R. & Jones, H. J. The role of mouse tumour models in the discovery and development of anticancer drugs. *Br. J. Cancer* **121**, 101–108 (2019).

44.   Paul, M. Training data and fine-tuned models. https://doi.org/10.5281/zenodo.14832406.

45.   Paul, M. Code on GitHub. https://github.com/maartenpaul/spheroidAnalysis; https://doi.org/10.5281/zenodo.15239560.

Fig. 1: Treatment response in cisplatin-sensitive and -resistant breast cancer PDX tumor spheroids. a, Schematic overview of spheroid generation from breast cancer patient-derived xenograft (PDX) models. **b,** Timeline of spheroid formation, treatment and sample processing. **c,** Maximum projection images from cisplatin-sensitive (CS) PDX spheroids either untreated (Unt) or treated with 10 µM Olaparib (Ola), 10 µM cisplatin (Cis-Pt) or 10 Gy radiation (IR) stained for cell apoptosis (cleaved caspase-3, red), cell proliferation (EdU, green) and nuclei (DAPI, blue). Scale bar: 100 µm. **d,** Maximum projection images from cisplatin-resistant (CR) PDX spheroids stained for cell apoptosis, proliferation and nuclei. Scale bar: 100 µm. **e,** Fold mean intensity of cleaved caspase-3 in CS tumor spheroids. Each point represents the fold change from the average intensity calculated from five separate Z-slices from five spheroids per tumor (n=4). **f,** Percentage of EdU+ cells in CS tumor spheroids. Each point represents a tumor (biological replicate; n=4). **g,** Fold mean intensity of cleaved caspase-3 in CR tumor spheroids. Each point represents the fold change from the average intensity calculated from five separate Z-slices from five spheroids per tumor (n=4). **h,** Percentage of EdU+ cells in CR tumor spheroids. Each point represents a tumor (biological replicate; n=4). **e-h,** Groups were compared using unpaired two-tailed Student's T-tests. **i,** Confocal images of CS tumor spheroids either untreated or treated with Olaparib, cisplatin or radiation (day 7). Spheroids were stained for 53BP1 (red) and nuclear stain (DAPI, blue). **j,** Percentage of cells with more than five 53BP1 foci in CS spheroids untreated or treated with Olaparib, cisplatin or radiation. Each point represents the percentage from one spheroid from which five separate Z-slices were analyzed for the four tumor models. Groups were compared using unpaired two-tailed Student's T-tests. **e-j,** Data points are shown with median.

Fig. 2: Automated spheroid recognition and size determination using AI. a, Schematic overview of AI-guided automated spheroid recognition pipeline. **b,** Heatmap of F1 scores calculated from separate CS training datasets or a combined training dataset. **c,** Percentage of correctly segmented spheroids when trained on itself (STS), on other single models or combined. Each point represents a tumor (biological replicate; n=4). Error bars indicate mean with standard deviation (SD). Groups were compared using an unpaired two-tailed Student's t-test with Welch's correction. **d,** Heatmap of F1 scores calculated from spheroid sets at 0, 4 and 7 days post-treatment using the combined CS training dataset. **e,** Comparison between manually-measured CS1 spheroid diameters (mm) and measured using automated recognition. r value shows Pearson correlation. Each data point represents a spheroid. **f,** Overlay of fold change

spheroid surface areas (A) manually calculated and determined with AI at 0, 4, and 7 days post-treatment. Each data point represents a spheroid. **g,** Pearson correlation scores when comparing spheroid diameters from manually-measured and automated spheroid recognition method, trained on CS1, CS2A, CS2B or combined data. **h,** Spheroid roundness measured by the ratio between height and width measured using automated recognition. Each data point represents a spheroid. **i,** Comparison between pixel-based and circular area of segmented CS1 spheroids. Each data point represents a spheroid. **j,** Pearson correlation scores when comparing measured and circular area from segmented spheroids. **e-f, h-i,** Exact numbers of spheroids for each condition can be found in Supplementary Data 2.

Fig. 3: Comparison between treatment response of CS and CR tumors. a-b, Bright-field images of CS1 (**a**) and CR1 (**b**) spheroids either untreated or treated with 10 µM Olaparib, 10 µM cisplatin or 10 Gy radiation. Scale bar: 100 µm **c-j,** Comparison between untreated (violin plot) and treated (points). Each data point represents a spheroid. Exact numbers of spheroids for each condition can be found in Supplementary Data 2. Groups were compared using parametric two-tailed Student's t-test with Welch's correction. CS1 (**c**), CS2A (**d**), CS2B (**e**) and CS3 (**f**), CR1 (**g**), CR2A (**h**), CR2B (**i**) and CR3 (**j**) spheroid surface areas at day 7.
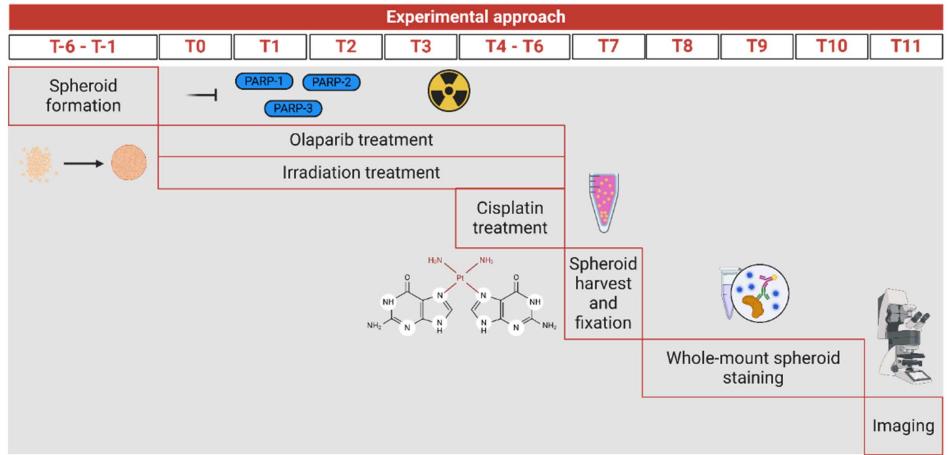
Fig. 4: Discrimination between treatment-sensitive and resistant tumors. a, Fold change of measured spheroid areas for CS1 and CR1 tumor models. Each data point represents a spheroid. **b,** Frequency distribution of measured spheroid areas from tumors treated with Olaparib. **c,** Frequency distribution of spheroid areas from tumors treated with cisplatin. **d-e,** Fold change of measured spheroid areas from CS (**d**) and CR(**e**) tumors. Data points are shown with median. Each data point represents a spheroid. **f,** Frequency distribution of spheroid areas from spheroids treated with irradiation. **g-h,** Percentage of spheroid data points measured below the calculated threshold for CS (**g**) and CR (**h**) tumor models. **a, d-e,** Exact numbers of spheroids for each condition can be found in Supplementary Data 2.
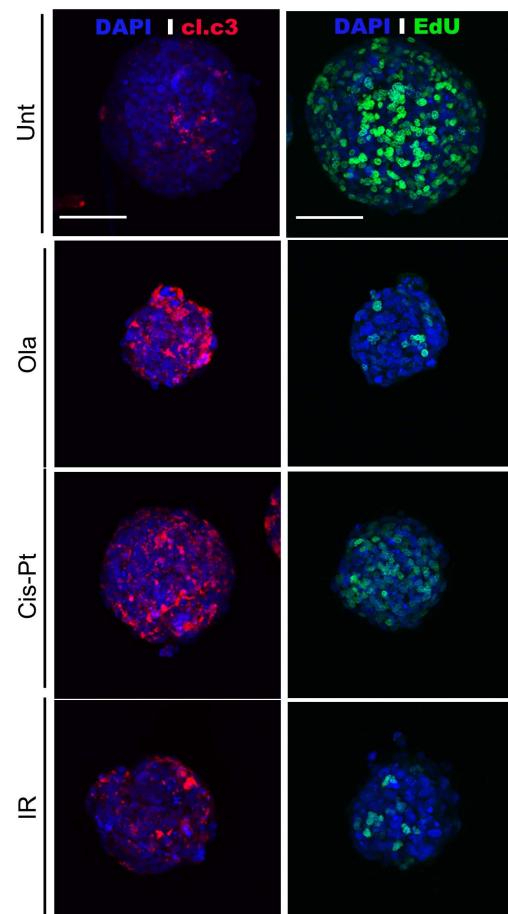
Editorial Summary:

Haspels et al. developed a high-throughput assay facilitated by automatic spheroid segmentation using deep learning. Measured differences in treatment response between cisplatin-sensitive and resistant tumors faithfully correspond with expected in vivo responses and the assay is able to discriminate between olaparib-sensitive and resistant tumors.
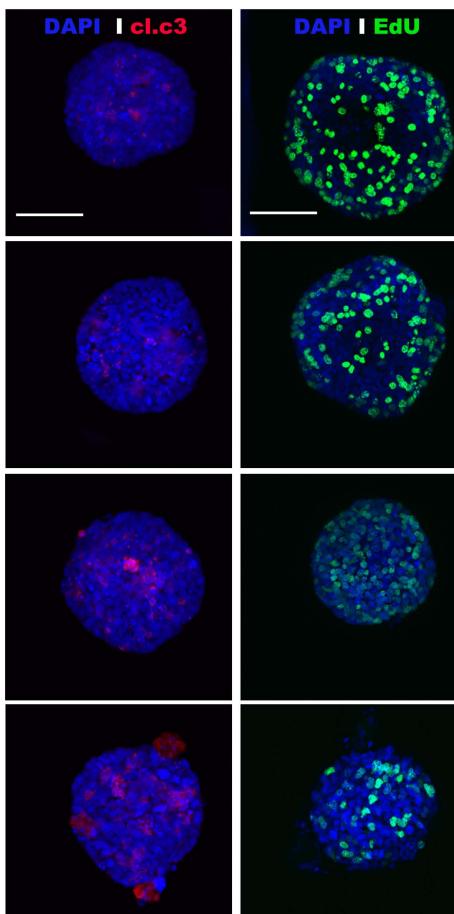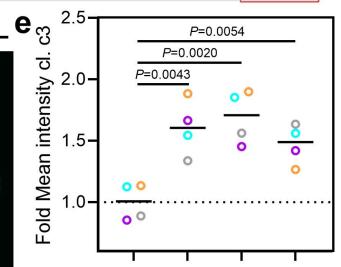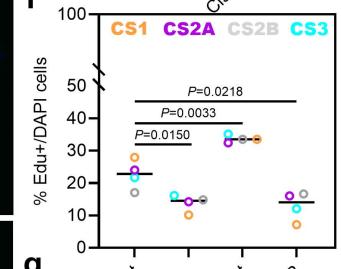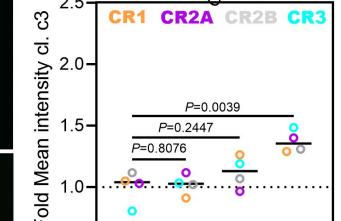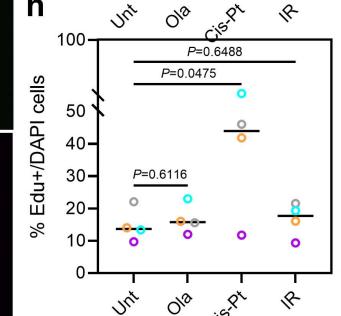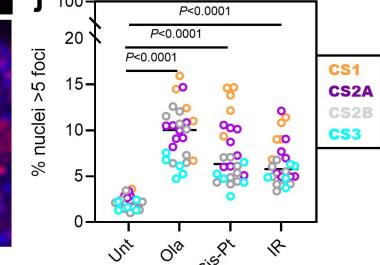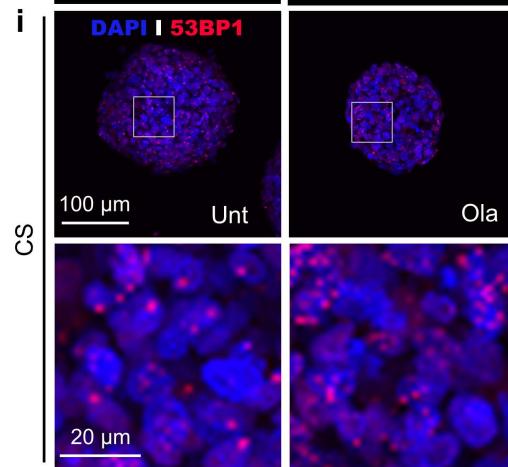
Peer review information:

Communications Medicine thanks Brian Karlberg, Evan F. Cromwel, Elad Katz, Zixuan Zhao and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.
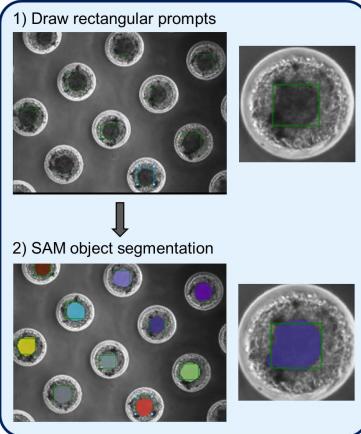
a

b Experimental approach

| T-6 – T-1 | T0 | T1 | T2 | T3 | T4 – T6 | T7 | T8 | T9 | T10 | T11 |

Spheroid formation

Olaparib treatment

Irradiation treatment

Cisplatin treatment

Spheroid harvest and fixation

Whole-mount spheroid staining

Imaging
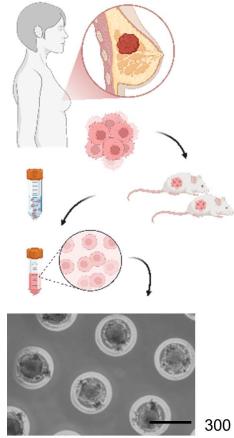
c CS PDX tumor

d CR PDX tumor

e CS PDX tumor

Fold Mean intensity cl. c3

P=0.0054
P=0.0020
P=0.0043

Unt   Ola   Cis-Pt   IR

f CS1   CS2A   CS2B   CS3

% Edu+/DAPI cells

P=0.0218
P=0.0033
P=0.0150

Unt   Ola   Cis-Pt   IR

g CR PDX tumor

CR1   CR2A   CR2B   CR3

Fold Mean intensity cl. c3

P=0.0039
P=0.2447
P=0.8076

Unt   Ola   Cis-Pt   IR

h

% Edu+/DAPI cells

P=0.6488
P=0.0475
P=0.6116

Unt   Ola   Cis-Pt   IR

i DAPI | 53BP1

CS   Unt   Ola   Cis-Pt   IR

j 53BP1 foci quant

% nuclei >5 foci

P<0.0001
P<0.0001
P<0.0001

CS1   CS2A   CS2B   CS3

Unt   Ola   Cis-Pt   IR

a

Semi-automated annotation
Napari - micro-SAM Annotator

1) Draw rectangular prompts

2) SAM object segmentation

Model finetuning and automatic segmentation
Jupyter notebook - micro-SAM

1) Finetune SAM model(s)
 - 1 image per timepoint/condition
 - 48 images total for training and validation each
 - 100 epochs

2) Run automatic instance segmentation

Input image   Automatic Instance Segmentation

3) Extract spheroid measurements
 - Area
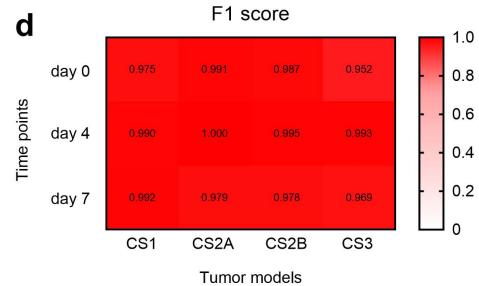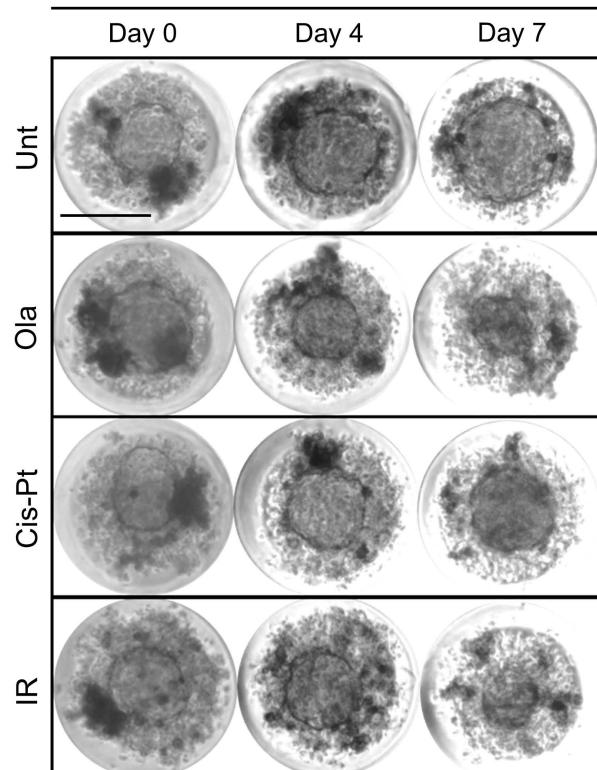 - Width/Height
 - Shape
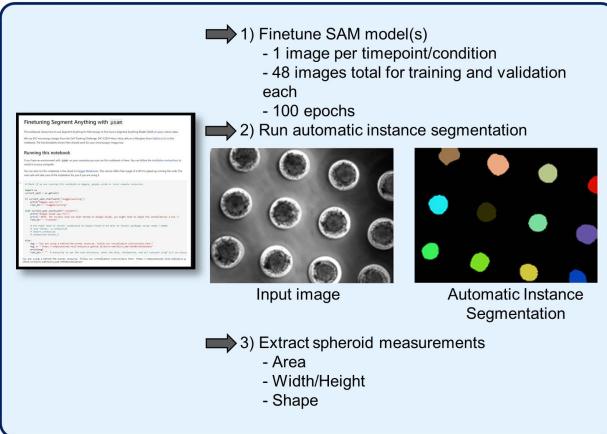
Segmentation Quality Assessment

1) Compare different training input data sets

Training input	Evaluation dataset
CS1	CS1
CS2A	CS2A
CS2B	CS2B
CS3	CS3
CR1	CR1
CR2A	CR2A
CR2B	CR2B
CR3	CR3

2) Compare with manual measurements

b F1 score

Training sets / Tumor Models

	CS1	CS2A	CS2B	CS3
CS1	0.961	0.965	0.965	0.742
CS2A	0.790	0.964	0.943	0.935
CS2B	0.815	0.970	0.987	0.971
CS3	0.464	0.868	0.942	0.955
All	0.986	0.990	0.987	0.972

c

% Correctly segmented spheroids — *P*=0.1018

x-axis: STS, CS1, CS2A, CS2B, CS3, Combined

CS1, CS2A, CS2B, CS3

d F1 score

Time points / Tumor models

	CS1	CS2A	CS2B	CS3
day 0	0.975	0.991	0.987	0.952
day 4	0.990	1.000	0.995	0.993
day 7	0.992	0.979	0.978	0.969

e Diameter CS1

Unt, Ola, Cis-Pt, IR

AI diameter (mm) vs Manual diameter (mm)

r=0.7864

f Training set CS1

Manual / AI

Fold change measured A

Unt, Ola, Cis-Pt, IR

x-axis: d0 d4 d7 d0 d4 d7 d0 d4 d7 d0 d4 d7

g Pearson Correlation

Training sets / Tumor models

	CS1	CS2A	CS2B
CS1	0.814	0.597	0.527
CS2A	0.466	0.796	0.639
CS2B	0.454	0.626	0.628
All	0.786	0.749	0.639

h Spheroid roundness

Ratio height/width

x-axis: CS1, CS2A, CS2B, CS3

i Area CS1

Unt, Ola, Cis-Pt, IR

measured Area (pixels) vs circular Area (pixels)

r=0.9887

j Pearson Correlation

Training set / Tumor models

	CS1	CS2A	CS2B	CS3
CSall	0.989	0.993	0.993	0.989

a CS1

	Day 0	Day 4	Day 7
Unt			
Ola			
Cis-Pt			
IR			

b CR1

	Day 0	Day 4	Day 7
Unt			
Ola			
Cis-Pt			
IR			

c CS1 day 7

P<0.0001
P<0.0001

Fold change measured A

Ola Cis-Pt IR

d CS2A day 7

P<0.0001
P<0.0001

Fold change measured A

Ola Cis-Pt IR

e CS2B day 7

P<0.0001
P<0.0001

Fold change measured A

Ola Cis-Pt IR

f CS3 day 7

P=0.0005
P<0.0001

Fold change measured A

Ola Cis-Pt IR

g CR1 day 7

P<0.0001
P=0.0339

Fold change measured A

Ola Cis-Pt IR

h CR2A day 7

P<0.0001
P=0.0124

Fold change measured A

Ola Cis-Pt IR

i CR2B day 7

P=0.0002
P=0.5087

Fold change measured A

Ola Cis-Pt IR

j CR3 day 7

P=0.0013
P<0.0001

Fold change measured A

Ola Cis-Pt IR