Presence of allele frequency heterogeneity defined by ctDNA profiling predicts unfavorable overall survival of NSCLC

Zhichao Liu[1,2#], Zhanhong Xie[1#], Shen Zhao[3#], Dawei Ye[4#], Xiuyu Cai[5], Bo Cheng[1], Caichen Li[1], Shan Xiong[1], Jianfu Li[1], Hengrui Liang[1], Zisheng Chen[6], Peng Liang[1], Jun Liu[1], Jianxing He[1], Wenhua Liang[1]

[1]Department of Thoracic Surgery and Oncology, the First Affiliated Hospital of Guangzhou Medical University, State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, Guangzhou Institute of Respiratory Health, Guangzhou 510120, China; [2]Department of Thoracic Surgery, Shanghai Chest Hospital, Shanghai Jiao Tong University, Shanghai 200030, China; [3]Department of Medical Oncology, Sun Yat-Sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou 510060, China; [4]Cancer Center, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China; [5]Department of General Internal Medicine, Sun Yat-Sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou 510060, China; [6]Department of Respiratory Medicine, the Sixth Affiliated Hospital of Guangzhou Medical University, Qingyuan People's Hospital, Qingyuan 511518, China

Contributions: (I) Conception and design: Z Liu, J Liu, J He, W Liang; (II) Administrative support: J He, W Liang, J Liu; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: B Cheng, C li, S Xiong, J Li, H Liang, Z Chen, P Liang; (V) Data analysis and interpretation: Z Liu, Z Xie, S Zhao, D Ye, C Xiu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.
[#]These authors contributed equally to this work.

Correspondence to: Jianxing He; Wenhua Liang; Jun Liu. Department of Thoracic Surgery and Oncology, The First Affiliated Hospital of Guangzhou Medical University, State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, Guangzhou Institute of Respiratory Health, No. 151, Yanjiang Rd, Guangzhou 510120, China. Email: drjianxing.he@gmail.com; liangwh1987@163.com; liujun9707@sina.com.

Background: The generation of subclonal (low-frequency) mutations is driven by tumor mutations and the relationship between the heterogeneity of tumor mutation abundance and non-small cell lung cancer (NSCLC) remains unknown. We investigate the role of allele frequency heterogeneity (AFH) defined by circulating tumor DNA (ctDNA) profiling in predicting prognosis in advanced NSCLC patients.

Methods: Publicly available data set of POPLAR (N=211) and OAK (N=642) trials were used for analyzing. A low ratio of allele frequency (AF) of a mutation to the maximum-somatic-allele-frequency (MSAF) was used to define the presence of AFH. The prognostic value of AF/MSAF ratio that was below a defined cutoff point in overall survival (OS) was evaluated using Cox-proportional hazards regression; and the structural break point was determined by LOESS regression and Chow test. The derived AFH was also explored in an independent cohort (N=259) of advanced NSCLC receiving first-line EGFR-TKIs from the First Affiliated Hospital of Guangzhou Medical University.

Results: In the POPLAR and OAK cohort, low AF/MSAF ratio was found to be significantly associated with unfavorable OS in univariate and multivariate analysis. The structural break point analysis demonstrated that AF/MSAF <10% could yield the optimal value to stratify patients with poor OS, which was applied for defining the presence of AFH. The presence of AFH significantly correlated with unfavorable OS in advanced NSCLC regardless of treatment arms (overall: HR 1.52, immunotherapy: HR 1.81, chemotherapy: HR 1.32, all P<0.05). In the exploratory EGFR-TKIs cohort, the presence of AFH was also significantly associated with shorter OS (HR 1.72, P=0.039).

Conclusions: Our results demonstrate that the presence of AFH predict unfavorable prognosis in advanced NSCLC despite which drug the patients used. The presence of AFH defined by ctDNA might provide an easily-accessible biomarker for risk stratification in the current clinical practice.

Keywords: Non-small cell lung cancer (NSCLC); heterogeneity; circulating tumor DNA (ctDNA)

Introduction

Recent studies have reported that the presence of subclonal mutations may compromise the anti-tumor immune response (1,2). These subclones may account for the genetic heterogeneity of tumor as well as resistance to treatment (3). As generation of subclones is driven by tumor mutations, the mutation distribution might serve as a surrogate measurement of subclones. Detection of mutations by circulating tumor DNA (ctDNA) has been increasingly accessible worldwide (4-7), and provides a comprehensive landscape across cancerous tissue for making it an ideal approach to assess the tumor heterogeneity. The variant allele fraction of different genetic mutations in ctDNA are often used to represent the clonal makeup of a tumor and high allele fraction mutations are more likely to be clonal, whereas lower values are more likely to be subclonal (8). Therefore, it is of great interest to explore the impact of distribution of mutation abundance on prognosis.

We hypothesized that the relatively low-frequency mutations in ctDNA might indicate allele frequency heterogeneity (AFH); a low ratio of allele frequency (AF) of a mutation to the maximum-somatic-allele-frequency (MSAF) was used to define the presence of AFH. The primary objective was, first, to explore whether AF/MSAF ratio that was below a defined cutoff point predicted prognosis of advanced NSCLC patients and, second, to evaluate the prognostic role of AFH (defined by the optimal AF/MSAF ratio) despite which drug the patients used.

Methods

Genomic and clinical data of advanced NSCLC patients from POPLAR (9) (N=211, NCT01903993) and OAK (10) (N=642, NCT02008227) trials from public database (https://doi.org/10.1038/s41591-018-0134-3) were included for primary analysis. Both POPLAR (randomized phase II trial) and OAK (randomized phase III trial) were designed to compared single agent atezolizumab with docetaxel as second/third-line therapy for patients with advanced NSCLC, and genomic alterations of pre-treatment plasma were estimated by FoundationOne CDx (F1CDx) as previously described (11,12).

The AF/MSAF status was categorized into two groups according to defined cutoff point with patients below and above the defined point; and the dividing points between groups were set from <1% to <50%. The prognostic value of AF/MSAF ratio that was below a defined cutoff point in overall survival (OS) was evaluated using Cox-proportional hazards regression. To find out the optimal prognostic cut-point for AF/MSAF, a series of AF/MSAF cutoffs were analyzed and the Wald-statistics of hazard ratios (HR) were fitted with a LOESS smoother (13). Then the structural break point of AF/MSAF was determined by Chow test (14,15), and was applied for separating patients into AFH present and AFH absent groups. The associations between presence of AFH and tumor size (SLD, sum of longest diameter of target lesions), or tumor-mutational burden (TMB) were also analyzed.

An independent cohort of 259 patients with advanced NSCLC receiving first-line EGFR-TKIs from the First Affiliated Hospital of Guangzhou Medical University (Jan 2012 to Dec 2018) was used for exploratory analysis to explore whether the prognostic role of AFH could be extended to different subsets of advanced NSCLC; genetic analysis of baseline plasma-derived ctDNA was confirmed by an targeted next generation sequencing panel (5) (Burning Rock Biotech, China) similar to the F1CDx test; ethical approval was obtained from institutional review boards and informed consent was waived due to the retrospective nature of this study. R software (Version 3.5.1) was used for all analyses and statistical significance was defined by a two-sided P<0.05.

Results

Development and analysis of allele frequency heterogeneity

This study used POPLAR and OAK trials of patients as primary cohort. A total of 853 advanced NSCLC patients that made up the blood biomarker-evaluable population were included for this analysis, with 784 (91.9%) EGFR-wild type patients.

In the univariable Cox proportional hazards regression model, low AF/MSAF ratio was found to be significantly associated with OS (*Figure 1A*). The cut-off valves (AF/
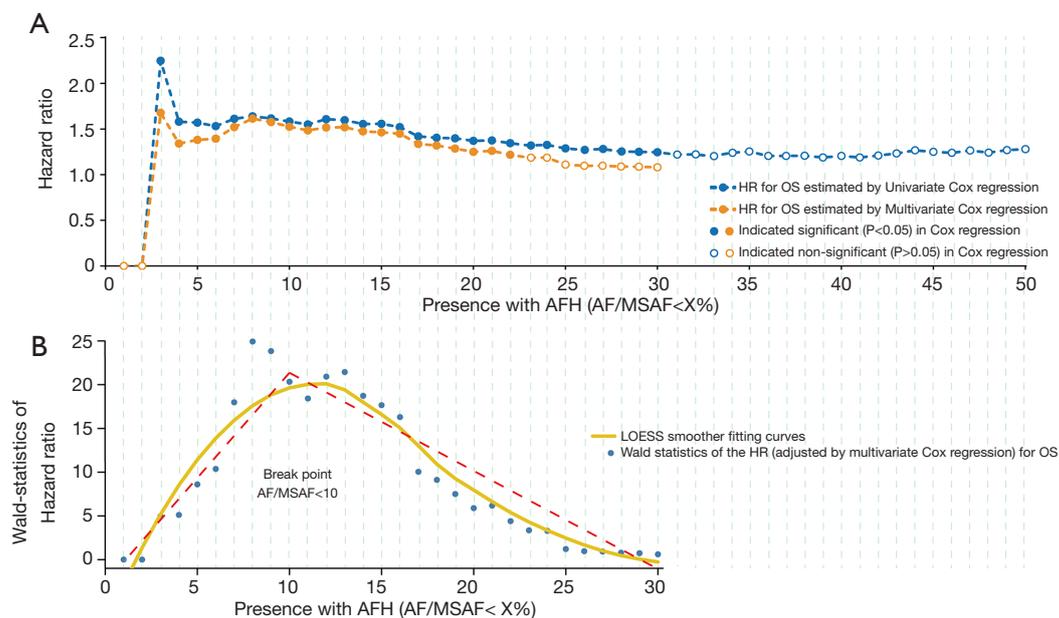
Figure 1 The associations between the AF/MSAF ratio that was below a defined cutoff point and overall survival. (A) Distribution of AF/MSAF cutoff and hazard ratio; (B) distribution of AF/MSAF cutoff and LOESS smoother fitting curves for Wald-statistics of hazard ratio, and determination of structural break points with use of the Chow test. We hypothesized that allele-frequency heterogeneity (AFH) of mutations might compromise the overall survival (OS) by its prognostic components (minor-subclones which represent tumor heterogeneity and only take up a small proportion of ctDNA). Based on the above hypothesis, the allele-frequency to the maximum-somatic-allele-frequency (AF/MSAF) ratio below a defined cutoff point is used as a marker of the presence of AFH. To explore the optimal cut-off value for defining the presence of AFH, a series of AF/MSAF ratios were tested and its impact on OS was analyzed; and the cut-off points of AF/MSAF of <1% to <50% were explored in the univariate analysis, then cut-off valves (AF/MSAF of <1% to <30%) with significant prognostic trends were selected into the multivariate analysis. For Figure B, the points indicate Wald-statistics of OS-HR for presence of AFH defined by different AF/MSAF cutoffs; a large Wald-statistic indicates a more statistically significant hazard ratio; structural break point analysis demonstrated that AF/MSAF with the break-point value (<10%) could yield the optimal predictive value for poor prognosis in advanced NSCLC; HRs and Wald-statistics were derived from the multivariate Cox-proportional hazard regression model. LOESS, locally estimated scatterplot smoothing; AF, allele frequency; MSAF, maximum somatic allele frequency; AFH, allele frequency heterogeneity. (The multivariate Cox-proportional hazards model included age, sex, race, histology, ECOG status, number of prior therapies, smoking status, SLD, driver mutations, the number of metastatic sites and TMB).

MSAF of <1% to <30%) with statistically significant prognostic trends were further examined in the multivariate analysis. In the multivariate Cox proportional hazards regression model that controlled all baseline factors (see more details in *Figure 1*), low AF/MSAF ratio (i.e., AF/MSAF <10%) was demonstrated as an independent prognostic predictor in patients with advanced NSCLC.

Figure 1B shows the LOESS fitting curves of Wald-statistics of OS-HR and AFH determined by different AF/MSAF cutoffs; and the corresponding structural break points for the AF/MSAF ratio. The structural break point analysis demonstrated that the presence of mutations with AF/MSAF <10% could yield the optimal predictive value to stratify

patients with poor OS, which was applied for defining the presence of AFH. Results (*Figure 2*) showed that both large tumor size (46.4% *vs.* 26.2%) and high TMB (68.6% *vs.* 24.8%) were significantly associated with the presence of AFH.

Prognostic significance of allele frequency heterogeneity

In multivariate survival analyses of the primary cohort (POPLAR and OAK), independent from TMB and SLD, the presence of AFH (AF/MSAF <10%) significantly correlated with unfavorable OS in advanced NSCLC (HR: 1.52, 95% CI: 1.27–1.83, P<0.001) (*Figure 3A*); and median OS was shorter in AFH present group than in AFH absent
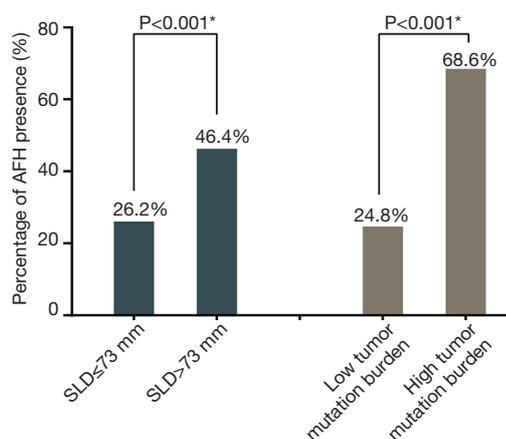
Figure 2 Distribution of the presence of AFH in SLD and TMB subgroups. AFH, allele frequency heterogeneity; SLD, sum of longest diameter of target lesions at baseline (dichotomized based on median value); TMB, tumor mutation burden [dichotomized based on high (≥16) vs. low (<16)]. *, statistical analysis performed by Pearson Chi-square test.

group (7.7 *vs.* 11.7 months). Additionally, the presence of AFH was significantly associated with shorter OS regardless of treatment arms (*Figure 3B,C*) [immunotherapy (HR: 1.81, 95% CI: 1.38–2.37, P<0.001; median OS: 7.9 *vs.* 15.6 months) and chemotherapy (HR: 1.32, 95% CI: 1.01–1.71, P=0.039; median OS: 7.3 *vs.* 9.7 months), compared with AFH absent group].

To unravel whether AFH defined by (AF/MSAF <10%) could identify patients with worse prognosis in different subsets of advanced NSCLC, another independent cohort of 259 patients with advanced NSCLC treated with EGFR-TKIs was used for exploratory analysis. In this exploratory cohort, results (*Figure 3D*) demonstrated that the presence of AFH was significantly associated with shorter OS (HR: 1.72, 95% CI: 1.02–2.91, P=0.039).

Discussion

Based on the impact of relatively low frequency mutations
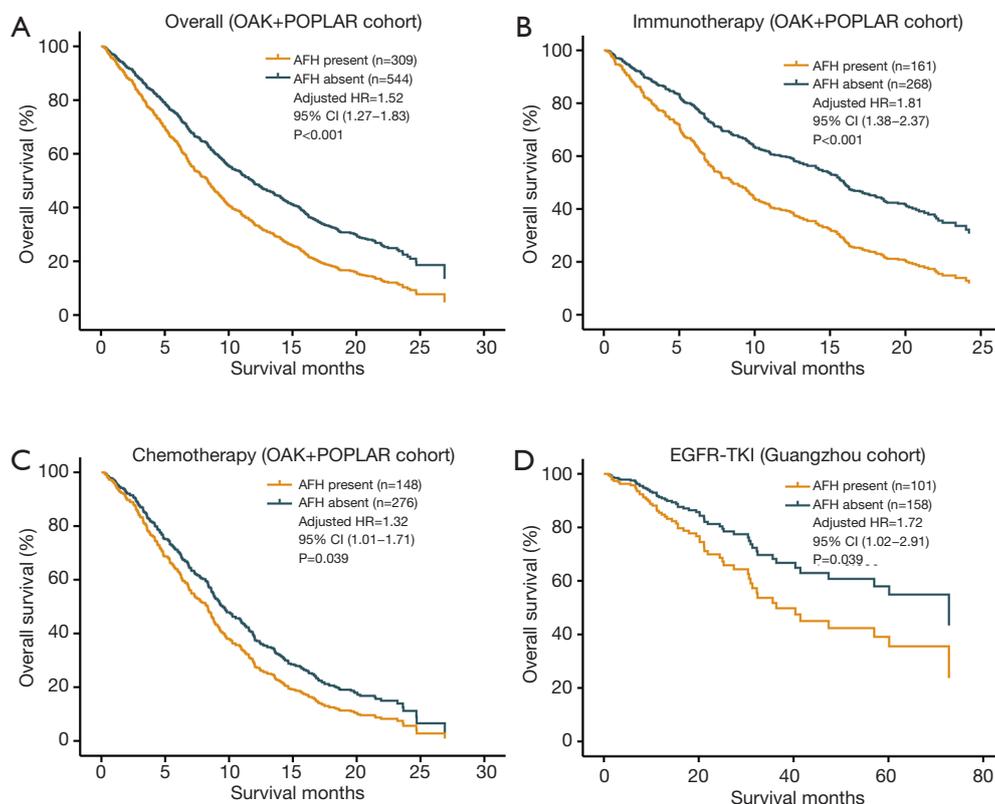

Figure 3 Multivariate survival analyses of AFH on overall survival in (A) the overall cohort of OAK and POPLAR, and (B) in the immunotherapy arm of OAK and POPLAR, and (C) in the chemotherapy arm of OAK and POPLAR, and (D) in the EGFR-TKI treatment cohort of Guangzhou, China. The multivariate Cox-proportional hazards model included age, sex, race, histology, ECOG status, number of prior therapies, smoking status, SLD, driver mutations, the number of metastatic sites and TMB. AFH, allele frequency heterogeneity.

on survival, which is simultaneously determined by AF and MSAF, we determined the presence of AFH by the existence of certain mutations with AF/MSAF <10%. Low AF/MSAF ratio was found to be significantly associated with unfavorable survival after controlling tumor bulk and TMB, suggesting that AF/MSAF could be used as an independent factor for risk stratification of advanced NSCLC patients. Additionally, the presence of AFH (cases with mutation of AF/MSAF <10%) represented an adverse prognostic factor for OS across treatment arms (chemotherapy, immunotherapy or EGFR-TKIs).

The OAK&POPLAR cohort (predominantly EGFR-wild-type NSCLC) and EGFR-TKI cohort are two different subsets, they represent the main population in advanced NSCLC nowadays. The exploratory analysis further revealed poorer prognosis for AF/MSAF <10% in EGFR-TKI cohort, which gave important assumption that the AFH defined by AF/MSAF <10% might be a universal prognostic factor across different subsets of advanced NSCLC.

As an increasing proportion of patients undergo ctDNA profiling to search for targeted drugs or to calculate TMB during the course of diagnosis or treatment. Importantly, ctDNA has proven to be a potentially source for tumor burden monitoring and treatment outcome monitoring (16-20). This study initially links relatively low frequency mutations (AFH) to prognostic outcomes by using accessible indicators (AF/MSAF). Low AF/MSAF ratio provides a way to mirror the presence of AFH and hence to identify patients with poor outcomes. It is difficult to obtain comprehensive evaluation information of tumor heterogeneity in current clinical practice, however, the presence of the AFH defined by low AF/MSAF ratio thus provides an easily-accessible biomarker for poor prognosis that will allow for improved risk stratification in patients with advanced NSCLC.

There are several limitations to this study. First, the clinical investigation was retrospective and the detection of ctDNA might be challenged by the technical approach. Second, we cannot report the mechanism responsible for the prognostic role of low AF/MSAF ratio. Third, there remains a lack of validation on the optimal cut point for AFH definition, though, it should be noted that, this is the first study to present real-world data concerning AFH defined by ctDNA and prognosis in advanced NSCLC; additionally, the significance of this study is to illustrate that the prognostic impact of AFH defined by relatively low frequency mutations might be a general character in all subsets of advanced NSCLC, rather than finding out an absolute definition for AFH. Finally, the AFH merits further validation and improvement.

In conclusion, presence of AFH (AF/MSAF <10%) is an independent predictive factor for poor prognosis in advanced NSCLC. By adding this simple calculation to ctDNA profiling, which incurs no further cost, we can evaluate or monitor tumor heterogeneity and its prognosis. Further experimental evaluation is still warranted to verify these findings.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. Ethical approval was obtained from institutional review boards and informed consent was waived due to the retrospective nature of this study.

References

1. McGranahan N, Furness AJ, Rosenthal R, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. Science 2016;351:1463-9.
2. Wolf Y, Bartok O, Patkar S, et al. UVB-Induced Tumor Heterogeneity Diminishes Immune Response in Melanoma. Cell 2019;179:219-235.e21.
3. Loeb LA. Human Cancers Express a Mutator Phenotype: Hypothesis, Origin, and Consequences. Cancer Res 2016;76:2057-9.
4. Murtaza M, Dawson SJ, Tsui DW, et al. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. Nature 2013;497:108-12.
5. Wang Z, Cheng Y, An T, et al. Detection of EGFR

mutations in plasma circulating tumour DNA as a selection criterion for first-line gefitinib treatment in patients with advanced lung adenocarcinoma (BENEFIT): a phase 2, single-arm, multicentre clinical trial. Lancet Respir Med 2018;6:681-90.

6. Yang XW, Xu W, Tian XY, et al. Diagnostic and prognostic value of KRAS mutations in circulating pancreatic ductal adenocarcinoma tumor DNA. Transl Cancer Res 2018;7:622-33.

7. Rathod A, Hopkins A, Rowland A, et al. Circulating cell free deoxyribonucleic acid for tracking early treatment response and disease progression in advanced cancers. Transl Cancer Res 2017;6:S1530-40.

8. Williams MJ, Werner B, Barnes CP, et al. Identification of neutral tumor evolution across cancer types. Nat Genet 2016;48:238-44.

9. Fehrenbacher L, Spira A, Ballinger M, 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 2016;387:1837-46.

10. Rittmeyer A, Barlesi F, Waterkamp D, et al. Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (OAK): a phase 3, open-label, multicentre randomised controlled trial. Lancet 2017;389:255-65.

11. Gandara DR, Paul SM, Kowanetz M, et al. Blood-based tumor mutational burden as a predictor of clinical benefit in non-small-cell lung cancer patients treated with atezolizumab. Nat Med 2018;24:1441-8.

12. Chalmers ZR, Connelly CF, Fabrizio D, et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. Genome Med 2017;9:34.

13. Cleveland WS, Grosse E, Shyu WM. Chapter 8: Local regression models. In: Chambers JM, Hastie TJ. editors. Statistical Models in S. Wadsworth & Brooks/Cole Advanced Books & Software, 1992.

14. Liang W, He J, Shen Y, et al. Impact of Examined Lymph Node Count on Precise Staging and Long-Term Survival of Resected Non-Small-Cell Lung Cancer: A Population Study of the US SEER Database and a Chinese Multi-Institutional Registry. J Clin Oncol 2017;35:1162-70.

15. Zeileis A, Kleiber C, Krämer W, et al. Testing and dating of structural changes in practice. Comput Stat Data Anal 2003;44:109-23.

16. Dawson SJ, Tsui DW, Murtaza M, et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. N Engl J Med 2013;368:1199-209.

17. Imamura F, Uchida J, Kukita Y, et al. Monitoring of treatment responses and clonal evolution of tumor cells by circulating tumor DNA of heterogeneous mutant EGFR genes in lung cancer. Lung Cancer 2016;94:68-73.

18. Garcia-Saenz JA, Ayllon P, Laig M, et al. Tumor burden monitoring using cell-free tumor DNA could be limited by tumor heterogeneity in advanced breast cancer and should be evaluated together with radiographic imaging. BMC Cancer 2017;17:210.

19. Wang YX, Li L, Cohen JD, et al. Prognostic Potential of Circulating Tumor DNA Measurement in Postoperative Surveillance of Nonmetastatic Colorectal Cancer. JAMA Oncol 2019;5:1118-23.

20. McDonald BR, Contente-Cuomo T, Sammut SJ, et al. Personalized circulating tumor DNA analysis to detect residual disease after neoadjuvant therapy in breast cancer. Sci Transl Med 2019. doi: 10.1126/scitranslmed.aax7392.