Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Diana Danelia April 6, 2021

Phospho-SAMHD1 as a potential prognostic biomarker for triple negative breast cancer patients

by

Diana Danelia

David Yu
Adviser

Biology

David Yu

Adviser

Andreas Fritz

Committee Member

Eladio Abreu

Committee Member

2021

Phospho-SAMHD1 as a potential prognostic biomarker for triple negative breast cancer patients

By

Diana Danelia

David Yu

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Biology

2021

Abstract

Phospho-SAMHD1 as a potential prognostic biomarker for triple negative breast cancer patients
By Diana Danelia

Triple negative breast cancer (TNBC) consists of a heterogeneous group of tumors characterized by a lack of estrogen and progesterone receptors, as well as lack of expression of the human epidermal growth factor receptor 2 (HER2) gene. Due to TNBC's lack of drug-targetable receptors, treatment options are generally limited to chemotherapy and resection. Additionally, metastasis as well as survival after metastatic relapse has been found to be shorter compared to other breast cancer subtypes, making TNBC one of the more aggressive breast cancers. Because of this, understanding the molecular characteristics of TNBC and identifying biomarkers that could be used for prognostic information could significantly help improve patient care by allowing for tailored treatments that improve patient outcomes.

Many cancer cells exhibit impaired intracellular dNTP pool homeostasis, leading to rapid cellular proliferation, enhanced mutagenesis, and dysregulation of the cell cycle. Intracellular dNTP pools are regulated partly by sterile alpha motif and histidine-aspartate domain containing protein 1 (SAMHD1) which acts as a deoxyribonucleoside triphosphate (dNTP) triphosphohydrolase. SAMHD1 has been found to have a role in restricting human immunodeficiency virus type 1 (HIV-1) and js also dysregulated in Aicardi Goutières syndrome (AGS) and cancers such as chronic lymphocytic leukemia (CLL). SAMHD1 has also been shown to have a role in promoting homologous recombination (HR) mediated DNA double strand break (DSB) repair independent of its dNTPase activity. Collectively, given SAMHD1's importance across cancers, as well as its roles in maintaining genome integrity, this protein demonstrates potential as a prognostic biomarker for TNBC.

SAMHD1 activity is partly regulated in a phosphorylation dependent manner throughout the cell cycle. SAMHD1 is phosphorylated at the C-terminal Thr592 amino acid residue by cycling-dependent kinases 1 and 2 (CDKs) in preparation for S-phase DNA replication. This lowers the rate of dNTP hydrolysis, coinciding with an increase in intracellular dNTPs. We show that for TNBC patients, individuals with higher p-SAMHD1 expression demonstrate worse progression-free survival outcomes, likely due to more aggressive tumors as a result of dNTP pool imbalances from impaired SAMHD1 dNTPase activity from its phosphorylation.

Phospho-SAMHD1 as a potential prognostic biomarker for triple negative breast cancer patients

By

Diana Danelia

David Yu

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Biology

2021

Acknowledgements

I want to thank my committee members for guiding me through my thesis. I also want to thank Dr. David Yu for the opportunity to work in his lab, and the Yu lab as a whole for inspiring my love of research during my undergraduate experience. In addition, I want to thank Dr. Li and Dr. Switchenko who made significant contributions to this project.

Table of Contents

Chapter 1: Introduction

Triple-negative breast cancer (TNBC) is defined by its lack of estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor-2 (HER2) protein expression. Epidemiological studies have shown that TNBC is most common in premenopausal women under 40 years old (an age group that also accounts for approximately 15-20% of all breast cancer patients) (Kumar and Aggarwal, 2016). When compared to other breast cancer subtypes, TNBC patient survival tends to be shorter, with mortality approximately 40% within the first five years after diagnosis. Additionally, TNBC tends to be highly invasive, with around 46% of TNBC patients developing distant metastasis. Average relapse in non-TNBC patients is 35-67 months, while in TNBC patients this lessens to only 19-40 months. Within three months after recurrence, the mortality rate of TNBC patients can also be as high as 75% (Yin et al., 2020). Due to the lack of receptors in TNBC, it is not sensitive to endocrine or molecular targeted therapies, and because of this, resection, radiation, and chemotherapy are the main treatment options. Given the aggressive nature of TNBC compared to other breast cancer types, it is imperative to research possible TNBC biomarkers that can be used as prognostic tools for tailored treatment options.

Sterile Alpha Motif and Histidine-Aspartic acid domain containing protein 1 (SAMHD1) is a deoxyribonucleotide triphosphate (dNTP) triphosphohydrolase and has been found to have a role in restricting human immunodeficiency virus type 1 (HIV-1) and other viral infections by depleting dNTPs required for reverse transcription and replication (Daddacha et al., 2017). Mutations in *SAMHD1* were first reported among patients with Aicardi-Goutières Syndrome (AGS), a rare genetic neuro-immunological disorder (Arnold et al., 2015b; Crow et al., 2015). It's theorized that the *SAMHD1* mutations in these patients interrupt normal cellular nucleic acid metabolism, triggering a hyper-interferon response (Crow and Livingston, 2008). Patients with AGS exhibit hyperactive innate immune response in the absence of infection, interfering with brain development and leading to death at earlier ages (Arnold et al., 2015b; Crow et al., 2015). SAMHD1 has also been shown to have a role in promoting homologous recombination (HR) mediated DNA double strand break (DSB) repair independent of its dNTPase activity. Through its interaction with C-terminal binding protein 1 (CtBP1) interacting protein (CtIP) and MRE11, SAMHD1 acts as a critical regulator of DNA end resection, an essential step for activation of HR (Coquel et al., 2018; Daddacha et al., 2017). Additionally, SAMHD1 has been shown recently to play an integral part in non-homologous end joining (NHEJ), where its dNTPase activity has an essential role in preventing high processivity and aberrant DNA synthesis prior to end-joining (Akimova et al., 2021). Furthermore, the Catalogue of Somatic Mutations in Cancer (COSMIC) has recorded 164 unique SAMHD1 mutations found in samples obtained from various cancer tissues, demonstrating its significance across cancers (Mauney and Hollis, 2018). Overall, SAMHD1 plays an important role in maintaining genome integrity and regulating dNTP pools within cells at appropriate levels for replication and repair, while keeping them below potentially mutagenic levels.

SAMHD1 contains an N-terminal sterile alpha motif (SAM) and histidine-aspartic acid containing domain (HD), with the SAM domain commonly being involved in protein-protein and

protein-DNA/RNA interactions. Meanwhile, the HD domain contains the dNTPase active site, regulatory sites, and the interfaces for enzyme oligomerization (Ji et al., 2014). SAMHD1 becomes catalytically active when it forms a homotetramer, with the HD domain tetramerizing in a nucleotide-dependent manner. In the absence of dNTPs, SAMHD1 exists in a monomer-dimer equilibrium regardless of phosphorylation state, however, as dNTP levels increase, such as in cycling cells, GTPs and dNTPs bind SAMHD1's allosteric sites, R1 and R2, resulting in tetramerization and protein activation (Arnold et al., 2015a; Morris and Taylor, 2019). SAMHD1 activity is regulated throughout the cell cycle in a phosphorylation dependent manner, where when a cell enters S-phase, SAMHD1 is phosphorylated at the C-terminal Thr592 residue by cycling-dependent kinases 1 and 2 (CDKs). This lowers the rate of dNTP hydrolysis, coinciding with an increase in intracellular dNTPs prior to S-phase DNA replication. By the end of M-phase, SAMHD1 dNTPase activity is restored through dephosphorylation by phosphatase PP2A-B55α, such that by the time cells reside in the noncycling G_0/quiescent state, the dephosphorylated SAMHD1 species predominates and correlates with lower dNTP pool levels (Mauney and Hollis, 2018; Morris and Taylor, 2019). Both phosphorylated and nonphosphorylated SAMHD1 are able to tetramerize, however, the stability of the tetramer varies between phosphorylation states. The CtD region of SAMHD1 is important for stabilizing the active tetramer structure of the protein and contains the site of Thr592 phosphorylation. Even though both p-SAMHD1 and SAMHD1 are able to tetramerize, phosphorylation of SAMHD1 disrupts Thr592 and Asp583 interactions within the CtD region, resulting in a more "open" and unstable tetramer conformation compared to dephosphorylated SAMHD1. As dNTP levels decrease, the unphosphorylated CtD region maintains tetramer stability, preventing the loss of activating dNTPs from SAMHD1's allosteric site, sustaining its dNTPase activity. In p-SAMHD1 however, dNTP activators are released from the allosteric site more easily, resulting in disassembly of the tetramer and overall impaired dNTPase activity (Arnold et al., 2015b). In summary, phosphorylation of SAMHD1 acts as a means of regulating SAMHD1 dNTPase activity, where p-SAMHD1 exhibits more unstable tetramers, reducing its regulation of dNTP pools.

Given SAMHD1's importance in genome integrity maintenance and cell-cycle dependent dNTP regulation, as well as its documented significance in multiple cancer types, uncovering its potential as a prognostic biomarker could significantly benefit TNBC patients who are currently limited in their treatment options. There is currently a deficit in studies demonstrating p-SAMHD1's potential clinical significance specifically for TNBC patients. This study aims to examine any correlations between p-SAMHD1 levels and patient outcomes in progression-free survival (PFS) and overall survival (OS). It's likely that high p-SAMHD1 expression correlates with worse patient outcomes, as SAMHD1 dNTPase activity will be impaired at high p-SAMHD1 levels, resulting in disruptions in dNTP pool homeostasis. Finding if p-SAMHD1 expression can be used as a prognostic factor for TNBC patients can help with further tailoring treatment options for patients to allow for more effective treatment response.

Chapter 2: Methods

Patient selection:

The patient cohort was selected from a maintained IRB approved database of patients diagnosed with triple negative breast cancer between 2002 and 2013 at Emory University's Winship Cancer Institute. A total of 87 TNBC patients from the database with sufficient tissue sample for corresponding sample staining were selected for analysis. Patient demographic, pathologic, and treatment outcome data were obtained from medical records.

Tissue microarray development:

Tissue microarrays (TMAs) were used for this study in order to allow for analysis of multiple patients on a single slide, reducing time and cost of analysis. Formalin-fixed paraffin embedded (FFPE) tissue blocks containing tumor specimens from each patient were acquired for TMA construction. Cores from each block were removed and placed into a TMA block in a specific row and column, and the coordinates were recorded to allow for accurate identification between tissue sample and corresponding clinical data. Three cores of normal tissue were used as controls to assess the efficacy of the staining, as well as to allow for core comparison on the same or separate TMAs to ensure quality.

Tissue microarray staining:

Once the TMAs were prepared, they were cut and stained for immunohistochemistry (IHC) staining with a mouse monoclonal P-SAMHD1 antibody from Abcam at a dilution of 1:600. TMAs underwent antigen retrieval using either Target Retrieval Solution (TRS) or Trilogy systems. The tissue blocks were subsequently blocked with 3% hydrogen peroxide for five minutes and then incubated with the primary antibody for 40 minutes. Dako's EnVision+ Dual Link System-horseradish peroxidase (HRP) was used for antibody detection.

IHC scoring:

After staining, the TMA sections were analyzed by a board-certified pathologist blinded to patient outcomes. Each sample was analyzed for intensity and percent spread of the staining. A previously validated scoring system was used to combine staining intensity and extent of staining into a quantified IHC score. Intensity of staining was scored from 1 to 3 and the percent spread was initially scored from 1-100%. Percent spread was then converted to a number between 1 and 3 with 1=0-15%, 2=51-80%, and 3=81-100%. The total IHC score was then calculated with the following formula: IHC= [(1+intensity)/3] x extent. Expression of p-SAMHD1 was evaluated using median values where an IHC score greater than 1.33 was generally defined as “high expression” and “low expression” as an IHC score less than 1.33.

Statistical analysis:

Kaplan-Meier survival analysis of p-SAMHD1 expression levels and PFS and OS outcomes were performed for all 87 patients. Analysis was performed using median values, with the cutoff for low expression being <= 1.33 and high expression being >1.33. After initial data for p-SAMHD1 expression and survival data was obtained, univariate analysis for OS, PFS, and other clinical pathological information associated with worse outcomes in TNBC patients was conducted using Cox proportional hazards models. Of the significant variables from the univariate analysis, further multivariate analysis with p-SAMHD1 expression was performed.

Chapter 3: Results

Descriptive statistics:

Of the patients in the study, 66.7% (n=54) were black and 23% (n=23) were white. 43.2% (n=35) had stage I cancer, 45.7% (n=37) stage II, and 11.1% (n=9) stage III. Within this cohort 63% (n=51) exhibited no lymph-vascular space invasion (LVSI) while 37% (n=30) did. The median age was 56.9 percent, with the youngest age group being 23 and the oldest being 83. According to the previously defined scoring definition of high and low expression, 65.5% (n=57) of patients exhibited low p-SAMHD1 expression (<= 1.33) and 34.5% (n=30) exhibited high expression (>1.33) (Table 1).

Survival Analysis: All patients

Kaplan-Meier survival analysis was performed for overall survival (OS) and progression free survival (PFS) across all 87 patients to examine any potential initial correlations between p-SAMHD1 expression and disease outcomes. PFS is defined as the period from treatment initiation to disease progression or worsening, while OS is the duration of patient survival from the start of treatment. OS is often used as a direct measure of clinical benefit, while PFS can be used as a direct or surrogate measure of clinical benefit, depending on the disease or response being observed (Hess et al., 2019). Analysis for OS showed no significant correlation between p-SAMHD1 expression for overall survival (OS), however the data still trended towards lower p-SAMHD1 expressing patients as having improved OS outcomes (figure 1a). P-SAMHD1 levels were significantly correlated with progression free survival (PFS), where low p-SAMHD1 levels was associated with improved outcomes, this time with statistical significance (figure 2a).

Cox proportional hazards: Univariate analysis

Univariate analysis using Cox proportional hazards models were performed for PFS and OS, as well as for other clinical pathological information associated with differences in TNBC patient outcomes such as stage, LVSI, and age (Table 2 and 3a). This was performed in order to assess for any confounding variables known to be associated with worse overall patient outcomes. Of these variables, stage and LVSI proved to be statistically significant for PFS for

both the hazard ratio (HR) p-value and the log-rank p-value, while for OS all variables remained statistically not significant. This indicates that for the dataset used in this study, LVSI and stage were two variables that were also significantly correlated with worse PFS outcomes among this cohort of TNBC patients. P-SAMHD1 was significantly correlated with worse PFS outcomes when performing the Cox proportional hazards models, where worse PFS outcomes were found for high p-SAMHD1 expression.

Cox proportional hazards: Multivariate analysis

Multivariate analysis was performed for p-SAMHD1, stage, and LVSI in order to examine p-SAMHD1 expression and PFS outcomes when accounting for LVSI and stage (Table 3b). None of the variables proved to be statistically significant, however higher p-SAMHD1 levels still showed a trend towards worse PFS outcomes.

Chapter 4: Discussion

Our study aimed to examine the clinical manifestations of previously established *in vitro* roles of SAMHD1 for TNBC patients. While there is data implicating SAMHD1's affects in multiple cancer types, this study explicitly links SAMHD1 and its phosphorylation-dependent regulation with TNBC patient PFS outcomes. Our data demonstrates that PFS outcomes among TNBC patients significantly differs depending on p-SAMHD1 expression, where increased p-SAMHD1 levels correlate with worse patient PFS outcomes when performing Kaplan Meier survival analysis (figure 1b). While p-SAMHD1 expression was not significantly correlated with OS outcomes, there was still a trend towards worse OS outcomes with high p-SAMHD1 expression in the Kaplan Meier survival analysis (figure 1a). When conducting multivariate analysis for LVSI, stage, and p-SAMHD1 for PFS, high p-SAMHD1 levels continue to trend toward worse PFS outcomes (table 3b). Because SAMHD1's dNTPase activity is dependent on its tetramerization and maintenance of the active state, it's likely that in patients with high expression of p-SAMH1, the phosphorylation prevents active tetramers of SAMHD1 persisting long enough to reduce dNTP levels, resulting in disruptions in dNTP pool homeostasis and subsequently manifesting in more aggressive tumors.

Increased dNTP levels have been shown to result in uncontrolled DNA replication and reduced replication fidelity that can contribute to cancer development. Many cancer types demonstrate impairment of dNTP pool homeostasis, resulting in their rapid cellular proliferation and enhanced mutagenesis (Kohnken et al., 2015). Precancerous cells are also characterized by increased mutagenesis, stimulation of genetic recombination, increased frequency of chromosomal abnormalities, DNA strand breaks, and cell death (Rentoft et al., 2016). While the mechanisms of mutagenesis from imbalanced dNTP pools are not fully understood, the symptoms of precancerous cells are consistent with what is observed with dNTP pool imbalance (Kohnken et al., 2015). dNTP pool abnormalities associated with SAMHD1 have been shown to lead to cancer development, where SAMHD1 mutations have been recognized as potential founding events in chronic lymphocytic leukemias (CLL) and in other leukemias as well as solid tumors (Franzolin et al., 2020). In a study by Rentoft et. Al. hemizygous deletions of SAMHD1 in mouse embryos led to an increase in dNTP pools similar to

that of tumors. Additionally, analysis of MMR-deficient yeast strains and human colorectal carcinoma cells showed that even small alterations in dNTP pools can result in a multiplicative increase in mutation rates (Rentoft et al., 2016). It is likely that increased p-SAMHD1 levels among TNBC patients result in increased dNTP pools, leading to dysequilibrium in cell replication and DNA repair processes, such as increased nucleotide mis-insertion, indirect inhibition of proofreading, or forced frameshift mutations which overall contribute to rapid uncontrolled cellular replication (Kohnken et al., 2015; Rentoft et al., 2016). In summary, in patients with high p-SAMHD1 levels, it's likely that the destabilized homotetramer structure increases dNTP pools, subsequently promoting enhanced mutagenesis and cell proliferation, resulting in more aggressive tumors that worsen patient PFS outcomes.

Collectively, this study suggests that increased dNTP levels due to SAMHD1 phosphorylation at Thr592 correlates with worse PFS outcomes among TNBC patients. This demonstrates SAMHD1's potential as a prognostic factor, where analysis of p-SAMHD1 levels in TNBC patients can be used to predict patient outcomes and allow for tailored treatment options. Despite the potential for p-SAMHD1 as a prognostic biomarker for TNBC, due to the nature of the limitations of this study, further research will be necessary. While we were able to see trends towards worse patient outcomes with high p-SAMHD1 expression, small patient numbers limited the sample size, restricting the types of analysis possible. Therefore, larger sample sizes in future studies would allow for deeper analysis (as well as validation of the trends from this study) to further elucidate p-SAMHD1's prognostic potential. Further clarity on the clinical impact of SAMHD1 will allow for further understanding of the implications of SAMHD1 expression for patients such as those with TNBC who are currently relatively limited in their treatment options.

Chapter 5: Figures

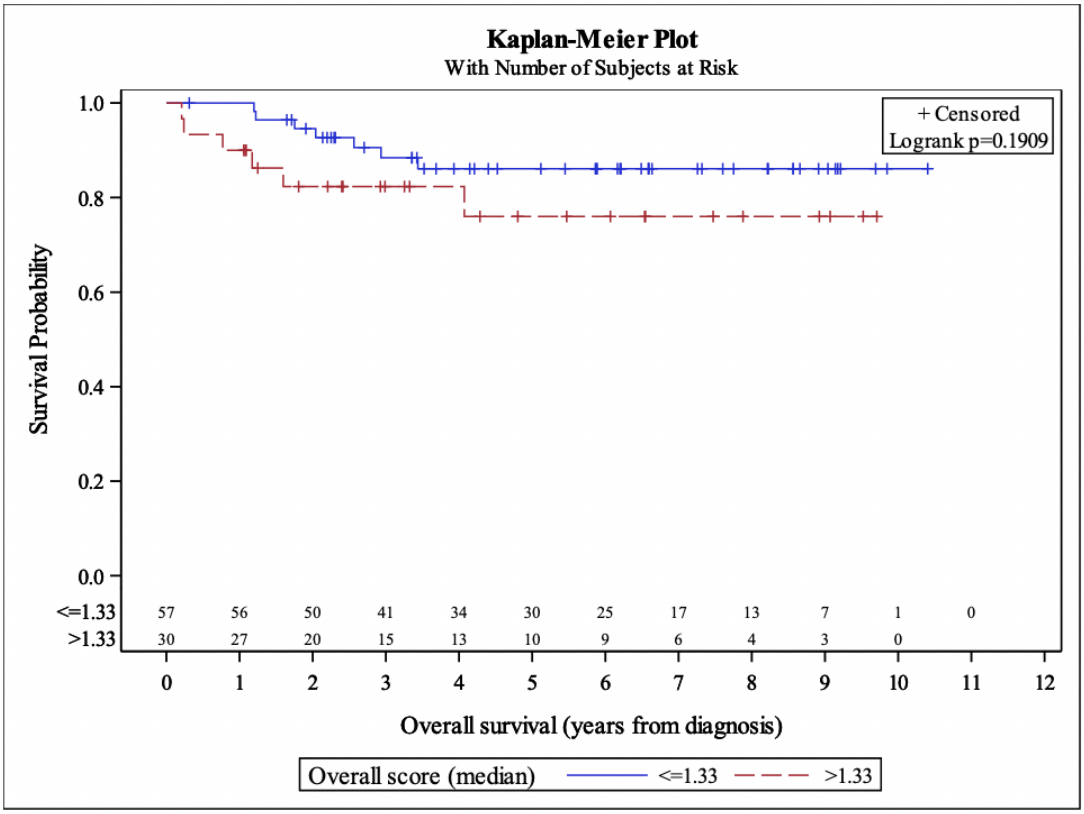
Table 1: Descriptive statistics

Variable	Level	N = 87	%
Overall score (median)	<=1.33	57	65.5
	>1.33	30	34.5
Overall score (quartile)	<=0.67	31	35.6
	>0.67, <=1.33	26	29.9
	>1.33, <=2.00	9	10.3
	>2.00	21	24.1
Race	Black	54	66.7
	White	23	28.4
	Other	4	4.9
	Missing	6	-
Stage	I	35	43.2
	II	37	45.7
	III	9	11.1
	Missing	6	-
Path T-stage	T1	42	51.9
	T2	33	40.7
	T3/4	6	7.4
	Missing	6	-

Variable	Level	N = 87	%
Grade	2	6	7.4
	3	75	92.6
	Missing	6	-
Surgery	Bilateral total mastectomy + ALND	6	7.4
	Bilateral total mastectomy + SLND	8	9.9
	Partial mastectomy	1	1.2
	Partial mastectomy + ALND	14	17.3
	Partial mastectomy + SLND	19	23.5
	R partial L total mastectomy + SLND	1	1.2
	Total mastectomy + ALND	21	25.9
	Total mastectomy + SLND	11	13.6
	Missing	6	-
Adjuvant therapy	No	81	100.0
	Missing	6	-
Margin status	Negative	80	98.8
	Positive	1	1.2
	Missing	6	-

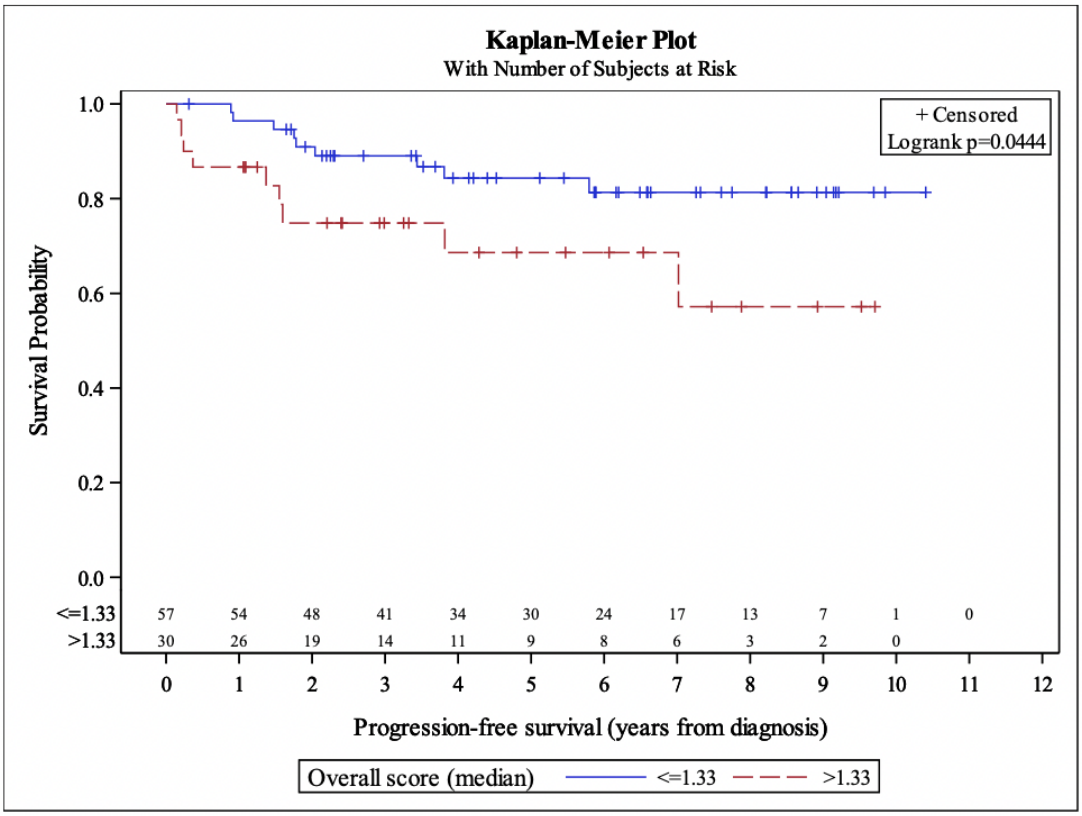
Variable	Level	N = 87	%
Margins-CIS	Negative	45	55.6
	N/A	36	44.4
	Missing	6	-
LVSI	No	51	63.0
	Yes	30	37.0
	Missing	6	-
Menopausal status	Premenopausal	14	18.7
	Perimenopausal	5	6.7
	Postmenopausal	56	74.7
	Missing	12	-
Overall score	Mean	1.72	-
	Median	1.33	-
	Minimum	0	-
	Maximum	6	-
	Std Dev	1.57	-
	Missing	0	-
Age	Mean	56.90	-
	Median	57	-
	Minimum	23	-
	Maximum	83	-
	Std Dev	12.95	-
	Missing	6	-

Figure 1a - OS - Median



Overall score (median)	No. of Subject	Event	Censored	Median Survival (95% CI)	1 Yr Survival	2 Yr Survival	5 Yr Survival
<=1.33	57	7 (12%)	50 (88%)	NA (NA, NA)	100.0% (NA, NA)	94.6% (84.1%, 98.2%)	86.1% (72.9%, 93.2%)
>1.33	30	6 (20%)	24 (80%)	NA (NA, NA)	90.0% (72.1%, 96.7%)	82.3% (62.5%, 92.3%)	76.0% (52.7%, 88.9%)

Figure 2a - PFS - Median



Overall score (median)	No. of Subject	Event	Censored	Median Survival (95% CI)	1 Yr Survival	2 Yr Survival	5 Yr Survival
<=1.33	57	9 (16%)	48 (84%)	NA (NA, NA)	96.4% (86.5%, 99.1%)	90.9% (79.6%, 96.1%)	84.3% (70.9%, 91.9%)
>1.33	30	9 (30%)	21 (70%)	NA (3.8, NA)	86.7% (68.3%, 94.8%)	74.8% (54.1%, 87.2%)	68.6% (45.6%, 83.5%)

Table 2: OS univariate analysis

			Overall survival (years from diagnosis)		

Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Log-rank P-value**
Overall score (median)	>1.33	30	2.04 (0.68-6.09)	0.200	0.191
	<=1.33	57	-	-	
Overall score (quartile)	>2.00	21	1.40 (0.31-6.26)	0.660	0.304
	>1.33, <=2.00	9	3.37 (0.75-15.10)	0.113	
	>0.67, <=1.33	26	0.93 (0.21-4.15)	0.922	
	<=0.67	31	-	-	
Race	Other	27	0.63 (0.17-2.34)	0.495	0.491
	Black	54	-	-	
Stage	III	9	7.96 (1.77-35.83)	**0.007**	**0.004**
	II	37	1.73 (0.41-7.26)	0.451	
	I	35	-	-	
Path T-stage	T3/4	6	7.03 (1.57-31.56)	**0.011**	**0.014**
	T2	33	1.67 (0.45-6.21)	0.448	
	T1	42	-	-	

			Overall survival (years from diagnosis)		
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Log-rank P-value**
Margins-CIS	N/A	36	1.53 (0.48-4.85)	0.469	0.466
	Negative	45	-	-	
LVSI	Yes	30	2.51 (0.79-7.90)	0.117	0.104
	No	51	-	-	
Menopausal status	Postmenopausal	56	0.92 (0.20-4.34)	0.917	0.902
	Perimenopausal	5	1.48 (0.13-16.36)	0.748	
	Premenopausal	14	-	-	
Overall score		87	1.09 (0.79-1.52)	0.592	-
Age		81	0.98 (0.93-1.03)	0.362	-

Table 3a: PFS univariate analysis

			Progression-free survival (years from diagnosis)		
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Log-rank P-value**
Overall score (median)	>1.33	30	2.51 (0.99-6.33)	0.052	**0.044**
	<=1.33	57	-	-	
Overall score (quartile)	>2.00	21	1.92 (0.62-5.97)	0.261	0.205
	>1.33, <=2.00	9	2.42 (0.60-9.69)	0.213	
	>0.67, <=1.33	26	0.60 (0.15-2.42)	0.477	
	<=0.67	31	-	-	
Race	Other	27	0.87 (0.31-2.48)	0.798	0.797
	Black	54	-	-	
Stage	III	9	4.94 (1.32-18.48)	**0.018**	**0.034**
	II	37	1.63 (0.53-5.01)	0.391	
	I	35	-	-	
Path T-stage	T3/4	6	3.65 (0.93-14.37)	0.064	0.140
	T2	33	1.28 (0.45-3.66)	0.647	
	T1	42	-	-	

			Progression-free survival (years from diagnosis)		
Covariate	**Level**	**N**	**Hazard Ratio (95% CI)**	**HR P-value**	**Log-rank P-value**
Margins-CIS	N/A	36	0.92 (0.35-2.42)	0.871	0.871
	Negative	45	-	-	
LVSI	Yes	30	2.75 (1.04-7.23)	**0.041**	**0.033**
	No	51	-	-	
Menopausal status	Postmenopausal	56	1.23 (0.27-5.54)	0.789	0.553
	Perimenopausal	5	2.60 (0.36-18.75)	0.342	
	Premenopausal	14	-	-	
Overall score		87	1.15 (0.88-1.51)	0.297	-
Age		81	1.00 (0.96-1.04)	0.931	-

Table 3b: PFS multivariate analysis

		Progression-free survival (years from diagnosis)	
Covariate	**Level**	**Hazard Ratio**	**HR P-value**
Overall score (median)	>1.33	1.93 (0.73-5.11)	0.183
	<=1.33	-	-
LVSI	Yes	2.01 (0.68-5.94)	0.209
	No	-	-
Stage	III	2.77 (0.62-12.28)	0.180
	II	1.15 (0.35-3.85)	0.818
	I	-	-

* Number of observations in the original data set = 87. Number of observations used = 81.

Chapter 6: References

1. Akimova, E., Gassner, F.J., Schubert, M., Rebhandl, S., Arzt, C., Rauscher, S., Tober, V., Zaborsky, N., Greil, R., and Geisberger, R. (2021). SAMHD1 restrains aberrant nucleotide insertions at repair junctions generated by DNA end joining. Nucleic Acids Res *49*, 2598-2608.
2. Arnold, L.H., Groom, H.C., Kunzelmann, S., Schwefel, D., Caswell, S.J., Ordonez, P., Mann, M.C., Rueschenbaum, S., Goldstone, D.C., Pennell, S*., et al.* (2015a). Phospho-dependent Regulation of SAMHD1 Oligomerisation Couples Catalysis and Restriction. PLoS Pathog *11*, e1005194.
3. Arnold, L.H., Groom, H.C.T., Kunzelmann, S., Schwefel, D., Caswell, S.J., Ordonez, P., Mann, M.C., Rueschenbaum, S., Goldstone, D.C., Pennell, S*., et al.* (2015b). Phospho-dependent Regulation of SAMHD1 Oligomerisation Couples Catalysis and Restriction. PLOS Pathogens *11*, e1005194.
4. Coquel, F., Silva, M.-J., Técher, H., Zadorozhny, K., Sharma, S., Nieminuszczy, J., Mettling, C., Dardillac, E., Barthe, A., Schmitz, A.-L*., et al.* (2018). SAMHD1 acts at stalled replication forks to prevent interferon induction. Nature *557*, 57-61.
5. Crow, Y.J., Chase, D.S., Lowenstein Schmidt, J., Szynkiewicz, M., Forte, G.M., Gornall, H.L., Oojageer, A., Anderson, B., Pizzino, A., Helman, G*., et al.* (2015). Characterization of human disease phenotypes associated with mutations in TREX1, RNASEH2A, RNASEH2B, RNASEH2C, SAMHD1, ADAR, and IFIH1. Am J Med Genet A *167a*, 296-312.
6. Crow, Y.J., and Livingston, J.H. (2008). Aicardi-Goutières syndrome: an important Mendelian mimic of congenital infection. Dev Med Child Neurol *50*, 410-416.
7. Daddacha, W., Koyen, A.E., Bastien, A.J., Head, P.E., Dhere, V.R., Nabeta, G.N., Connolly, E.C., Werner, E., Madden, M.Z., Daly, M.B*., et al.* (2017). SAMHD1 Promotes DNA End Resection to Facilitate DNA Repair by Homologous Recombination. Cell Rep *20*, 1921-1935.
8. Franzolin, E., Coletta, S., Ferraro, P., Pontarin, G., D'Aronco, G., Stevanoni, M., Palumbo, E., Cagnin, S., Bertoldi, L., Feltrin, E*., et al.* (2020). SAMHD1-deficient fibroblasts from Aicardi-Goutières Syndrome patients can escape senescence and accumulate mutations. Faseb j *34*, 631-647.
9. Hess, L.M., Brnabic, A., Mason, O., Lee, P., and Barker, S. (2019). Relationship between Progression-free Survival and Overall Survival in Randomized Clinical Trials of Targeted and Biologic Agents in Oncology. J Cancer *10*, 3717-3727.
10. Ji, X., Tang, C., Zhao, Q., Wang, W., and Xiong, Y. (2014). Structural basis of cellular dNTP regulation by SAMHD1. Proc Natl Acad Sci U S A *111*, E4305-4314.
11. Kohnken, R., Kodigepalli, K.M., and Wu, L. (2015). Regulation of deoxynucleotide metabolism in cancer: novel mechanisms and therapeutic implications. Molecular Cancer *14*, 176.
12. Kumar, P., and Aggarwal, R. (2016). An overview of triple-negative breast cancer. Arch Gynecol Obstet *293*, 247-269.
13. Mauney, C.H., and Hollis, T. (2018). SAMHD1: Recurring roles in cell cycle, viral restriction, cancer, and innate immunity. Autoimmunity *51*, 96-110.
14. Morris, E.R., and Taylor, I.A. (2019). The missing link: allostery and catalysis in the anti-viral protein SAMHD1. Biochemical Society Transactions *47*, 1013-1027.
15. Rentoft, M., Lindell, K., Tran, P., Chabes, A.L., Buckland, R.J., Watt, D.L., Marjavaara, L., Nilsson, A.K., Melin, B., Trygg, J*., et al.* (2016). Heterozygous colon cancer-associated mutations of SAMHD1 have functional significance. Proceedings of the National Academy of Sciences *113*, 4723.
16. Yin, L., Duan, J.J., Bian, X.W., and Yu, S.C. (2020). Triple-negative breast cancer molecular subtyping and treatment progress. Breast Cancer Res *22*, 61.