Genome-Wide Analysis and Characterization of an Online Sarcoma Cohort

K.E. Barnholt[1], C.B. Do[1], A.K. Kiefer[1], M. Nelson[1], J.E. Garber[2], M.E. Robson[3], R.G. Maki[4], G.D. Demetri[5], A. Wojcicki[1].

[1]23andMe, Inc, Mountain View, CA; [2]Dana-Farber Cancer Institute, Boston, MA; [3]Memorial Sloan-Kettering Cancer Center, New York, NY; [4]Mount Sinai School of Medicine, New York, NY; [5]Ludwig Center at Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA

Introduction

Sarcomas represent an uncommon, complex and heterogeneous group of mesenchymal-derived cancers, manifesting as over 50 distinct subtypes and striking just 14,000 individuals in the United States annually[1]. As such, recruitment of an adequately sized cohort to power sarcoma genome-wide discoveries presents a double challenge. Traditional barriers to participation include proximity of clinical and research centers of excellence, as well as the limits of patient motivation or ability to travel[2,3,4]. The web-based research platform offered by 23andMe provides increased accessibility to patient and family participation, facilitating rapid recruitment of a relevant population and enabling a large-scale genome-wide association study (GWAS) of rare diseases such as sarcoma.

Methods

Patients with any history of ever having had a diagnosis of sarcoma or mesenchymal neoplasm were recruited through web and email campaigns, patient advocacy groups, physician offices, and events. Participants provided IRB-approved consent, completed surveys, and received updates about research progress through an online account. In collaboration with an uncompensated panel of academic experts, an online survey was developed to collect patient-reported data on diagnosis, family history, symptoms and treatment. Association scans were conducted across a set of 1,047,958 SNPs, using 714 unrelated sarcoma cases of European ancestry and over 88,000 unrelated population controls from the 23andMe database.

Results

RECRUITMENT and ENGAGEMENT

In two years, 23andMe's web-based approach has led to accrual of one of the largest consented, genotyped, engaged, recontactable sarcoma cohort to date. Over 889 sarcoma patients have enrolled, 811 have been genotyped, and 668 have completed the sarcoma baseline survey.

- 85% of participants have been recruited by online means (33% by advocacy partner web campaigns/emails, 18% by 23andMe web campaigns/emails, 15% social networking sites, 13% email distribution lists, and 5% other internet search/blogs)
- High study completion rate and engagement: over 82% of genotyped sarcoma participants have taken the sarcoma survey; over 59% have completed more than five 23andMe online surveys

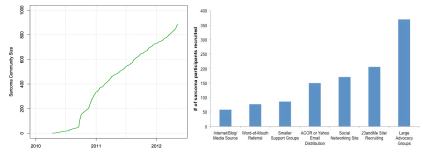

Figure 1. (A) Growth of sarcoma research community from launch in April 2010 to May 2012. Currently, over 889 people who have or who have had sarcoma are enrolled. The recruitment goal is 1,000 participants. (B) Recruitment channels for sarcoma cohort

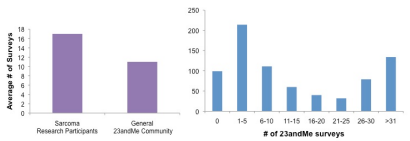

Figure 3. Survey participation by sarcoma research participants (A) as compared to customers in the general 23andMe database and (B) as a breakdown of total number of surveys completed. Sarcoma research participants are asked to take a sarcoma-specific survey and have the option to take any of the other 23andMe surveys covering a variety of topics. Survey completion is voluntary; no compensation is provided.

DEMOGRAPHICS

- The cohort is primarily of European ancestry (89%), disproportionately female (72%), with an average age of 50 (± 16) years. The majority of participants reside in the United States.
- Controls are also primarily of European ancestry (77%), 43.9% female, with an average age of 46 (± 17) years.

International Distribution of Sarcoma Community Members

Country	Sarcoma members	Percent of total
USA	720	82%
Great Britain	80	9%
Canada	34	4%
Australia	18	2%
Other	26	3%

Number of US Sarcoma Community Members by State

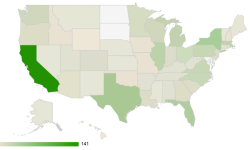

Figure 4. Global demographics of sarcoma participants, including 19 countries and 47 US states.

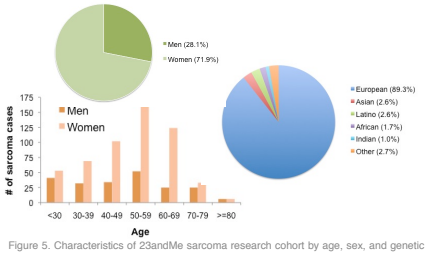

Figure 5. Characteristics of 23andMe sarcoma research cohort by age, sex, and genetic ancestry (N = 769 genotyped sarcoma cases).

DIAGNOSIS, SYMPTOMS, and TREATMENTS

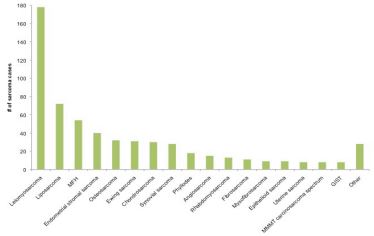

Figure 6. Breakdown of sarcoma subtypes in online cohort.

- Time since diagnosis - MEAN: 5.85 ± 6.42 years ago; MEDIAN: 4 years ago.
- Leiomyosarcoma (N = 178), liposarcoma (N = 72), and malignant fibrous histiocytoma/undifferentiated sarcoma (N = 54) are the most common subtypes represented in the online sarcoma cohort
- **Fifteen people report having a first degree relative with a sarcoma diagnosis (mother: 5, father: 5, sister: 3, brother: 1, daughter: 1).**
- 78% of participants experienced symptoms prior to diagnosis; the most common symptom experienced was pain
- Participants reported pain most often in the abdomen (34%), the leg (31%), and the pelvis (23%).
- Over 37% of participants report undergoing active treatment of some type.
- The top five treatments reported by participants to have been received for sarcoma therapy include: doxorubicin, gemcitabine + docetaxel, ifosfamide, doxorubicin + ifosfamide, and cisplatin.
- The most commonly reported side effects include hair loss, nausea, and low white blood cell count.

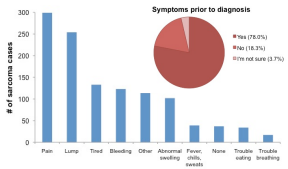

Figure 7. Sarcoma symptoms experienced at first diagnosis

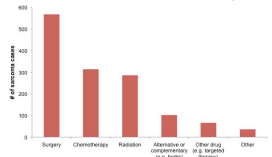

Figure 8. Distribution of treatments received by participants for sarcoma

GENOME WIDE ANALYSIS

Initial results have identified no significant genome-wide associations for general sarcoma risk, despite having >90% power to detect risk variants with >5% minor allele frequency and odds ratio >2.5, suggesting the absence of common variants with strong shared effects across sarcoma subtypes.

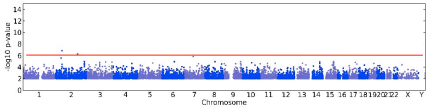

Figure 9. Manhattan plot of genome-wide association study (GWAS) results from sarcoma cohort. The red line represents the "suggestive" threshold. No significant hits.

Subtype	SNP	Region	Gene	P value	MAF	HWE
Any sarcoma	rs2715053	chr2:51556409	LOC730100/LOC100128229	1.3e-7	0.360	0.96
Any sarcoma	rs6725973	chr2:170660583	UBR3/MYO3B	5.3e-7	0.041	0.541
Uterine Leiomyosarcoma (UL)	rs2294998	chr20:60994578	DIDO1	3.4e-7	0.337	0.067
Uterine Leiomyosarcoma (UL)	rs7698736	chr4:142050902	RNF150	6.3e-7	0.376	0.029
Leiomyosarcoma + UL	rs7078493	chr10:98194085	TLL2	1.8e-7	0.004	1.000
Leiomyosarcoma + UL	rs12739933	chr1:53776336	GLIS1	6.8e-7	0.383	0.584
Leiomyosarcoma (no UL)	rs16935830	chr9:1913958	DMRT2/SMARCA2	6.2e-7	0.064	0.926
Osteosarcoma + chondrosarcoma	rs11765480	chr7:136622567	PTN	2.4e-7	0.373	0.725

Table 1. Suggestive associations (p < 1e-6) from all sarcoma and sarcoma subtype GWAS. N = 714 sarcoma; N = 36 uterine leiomyosarcoma; N = 169 leiomyosarcoma + uterine leiomyosarcoma; N = 133 leiomyosarcoma (no uterine leiomyosarcoma); N = 58 osteosarcoma + chondrosarcoma. N = > 88,000 controls for all GWAS. No suggestive associations from liposarcoma GWAS.

Discussion

This study demonstrates the feasibility of both rapid recruitment and longitudinal engagement of patients through a web-based research platform. A combination of online recruitment methods by sarcoma advocacy groups, 23andMe, and participants has been key to accruing one of the largest geographically and phenotypically diverse genotyped sarcoma cohorts in the world, and has enabled one of the first sizable genomic studies of this rare disease. Investigation of associations within genetically more homogeneous sarcoma subtypes remains a promising avenue for future exploration, although will require additional recruitment to achieve adequate statistical power. Web-based genetic research has the potential to change the one size fits all approach to clinical research and transform how sarcoma is diagnosed and treated.

Acknowledgments

We thank the study participants for their support and time spent contributing to research. We also thank the other 23andMe customers who have consented to participate in research. We also thank our advocacy partners who have been invaluable resources for this research, including Sarcoma Alliance, Beat Sarcoma, Sarcoma Foundation of America, Sarcoma UK, and Association of Cancer Online Resources (ACOR).

References

1. National Cancer Institute. Surveillance, Epidemiology, and End Results (SEER) Program and the National Center for Health Statistics. http://seer.cancer.gov/
2. Liddy Shriver Sarcoma Initiative. Clinical Trials and Travel. http://sarcomahelp.org/sarcoma_videos/gorlick_travel.html
3. J.M.Watson, D.J.Torgerson. Increasing recruitment to randomised trials: a review of randomised controlled trials. *BMC Medical Research Methodology*. 2006: 6(34).
4. National Cancer Institute. Cancer Clinical Trials: The In-Depth Program. 2002. Retrieved from http://www.cancer.gov/clinicaltrials/ learningabout/in-depth-program/