

Genetics Working Group

Updates

Current Priorities

Vision for next 5 years

Updates...

6/2015

Concept **Genomic alterations in radiation-related breast cancer using Array-CGH (Comparative Genomic Hybridization)**

PI **Rose Yang, NCI**

Aims Prevalence/ patterns of DNA copy number changes in radiation-exposed breast tumor tissues

RESEARCH ARTICLE

Characterization of Genomic Alterations in Radiation-Associated Breast Cancer among Childhood Cancer Survivors, Using Comparative Genomic Hybridization (CGH) Arrays

Xiaohong R. Yang^{1*}, J. Keith Killian², Sue Hammond³, Laura S. Burke¹, Hunter Bennett¹, Yonghong Wang², Sean R. Davis², Louise C. Strong⁴, Joseph Neglia⁵, Marilyn Stovall⁶, Rita E. Weathers⁶, Leslie L. Robison⁷, Smita Bhatia⁸, Kiyohiko Mabuchi¹, Peter D. Inskip¹, Paul Meltzer²



Array CGH data of second breast cancer after chest RT for HL (n=32)	CCSS	TCGA
ER +ve staining:	68%	65%
High proliferative index (Ki-67 staining):	77%	—
Amplifications of 17q12 region containing HER2	<u>40%</u>	<u>14%</u>
FGFR1 (drives endocrine Rx resistance)	<u>20%</u>	<u>12%</u>
CCND1 (oncogene – cell cycling regulatory protein)	<u>28%</u>	<u>14%</u>

Conclusions: Radiation-related breast cancers are enriched for an amplifier genomic subgroup with highly proliferative breast tumors

Concept

Telomere length and Second Malignancy in Pediatric Cancer Survivors

PI MM Gramatges, Baylor

Hypothesis Shortened germline telomere length plays a role in development of SMNs in childhood cancer survivors

Aims Investigate relation between telomere length in buccal DNA samples and SMNs

Methods Matched case-control study design
cases: n=147; controls: n=147
Breast ca (n=68); thyroid ca (n=48); sarcoma (n=31)
qPCR analysis to measure telomere length

Human Cancer Biology

See related article by Shay, p. 779

Telomere Content and Risk of Second Malignant Neoplasm in Survivors of Childhood Cancer: A Report from the Childhood Cancer Survivor Study

Maria M. Gramatges¹, Qi Liu³, Yutaka Yasui³, M. Fatih Okcu¹, Joseph P. Neglia⁴, Louise C. Strong², Gregory T. Armstrong⁵, Leslie L. Robison⁵, and Smita Bhatia⁶

Clin Cancer Res; 20(4); 904–11. ©2013 AACR.

Conclusions: A relation between lower telomere content and treatment-related thyroid cancer was observed, suggesting that shorter telomeres may contribute to certain SMNs in childhood cancer survivors.

Telomere homeostasis and thyroid cancer – Ancillary Study

R01 CA194473 (PI: MM Gramatges)

(PQB-1) Defects in telomere maintenance associated with thyroid SMNs in childhood cancer survivors

- Aim 1** Investigate the association between SNPs in telomere maintenance genes and thyroid SMN risk
- Aim 2** Characterize rare genetic telomere maintenance defects in survivors with thyroid SMN (targeted DNA sequencing, functional studies)
- Aim 3:** Extend and improve upon the previously validated risk prediction model for thyroid SMN.

Concept

**Genetic Alterations in Second Malignant
Neoplasms**

PI

Jean Nakamura, UCSF

Genetic Alterations in Second Malignant Neoplasms

- Aim 1** Determine LOH in tumor suppressor genes (previously identified in the *Nf1* mutant mouse model) in SMNs
SNP genotyping (Taqman) on DNA isolated from FFPE SMN samples
- Aim 2** Determine whether transcript levels of candidate tumor suppressor genes are reduced in SMNs
- Aim 3** Sequence exons of 360 frequently mutated genes in human cancers/genes in known cancer-relevant pathways.
- Specimens** Unstained, fixed sections of SMN samples; paired normal samples
- Funding** St. Baldrick's Foundation

SMN type	Requested SMN tissue/ paired germline DNA	Received SMN tissue/ paired germline DNA
Breast	48/48	43/39
Meningioma	23/23	6/0
CNS	18/18	2/0
Sarcoma	10/10	4/0

Genetic Alterations in Second Malignant Neoplasms

Status

- Isolated DNA from most of the CCSS tumor slides and ran initial QA assays (alongside DNA from freshly obtained SMN samples collected from UCSF)
 - Quality of FFPE samples to be significantly worse than the quality of frozen.
- Performed whole genome amplification of DNA from the CCSS samples,
 - Did not sufficiently raise the DNA quality metrics to be comparable to those of freshly isolated DNA.
- Performing targeted exome sequencing (not whole exome) – would yield information on far fewer genes
 - Compromise to obtain high quality sequencing on fewer genes (not mediocre sequencing across the exome).
- Sequencing frozen tumor samples to obtain initial characterizations
 - will use information on nature of genetic alterations to guide next steps with the CCSS samples.
- Seeking extramural funding
 - revising an R01 application that proposes to sequence the CCSS samples.

Concept

**Genetic Epidemiology of Basal Cell
Carcinoma in Childhood Cancer Survivors**

PI

Stella Davies, Cincinnati

Funding: NIH (U01)

Genetic Epidemiology of Basal Cell Carcinoma in Childhood Cancer Survivors

- Aims** Identify association between polymorphisms in genes related to radiation sensitivity and basal cell carcinoma
- Methods** Case-control study design (cases: n=321; controls: n=825)
Candidate gene approach (*128 SNPs in radiation-related DNA repair genes*)
- Findings** SNP rs1800058 in *ATM* gene was associated with increased risk of BCC (OR=1.84, p=0.02)
(*adjusting for primary cancer diagnosis, radiation dose, age at cancer diagnosis and time from diagnosis*)
- Status** Manuscript in preparation

Concept

Susceptibility genes for radiation-induced breast cancer after Hodgkin lymphoma

PI

van Leeuwen/ Netherlands

Radiation-induced breast cancer after HL – germline variants

- Aim** Examine gene-environment interactions in patients with radiation-related breast cancer after HL
- Design** Case-control study design
Cases: Caucasians with HL and secondary breast cancer
Controls: Caucasians with HL free of breast cancer until date of inclusion as controls
Matching criteria: age at dx of HL; calendar year of HL dx; length of f/u \geq cases; exposure to supra-diaphragmatic radiation
2nd comparison with available data on 3000 young BC without HL
- Platform** Illumina iSelect 200,000 targeted SNP chip
- Status** Cases (n=332)/ controls (n=608) identified [CCSS/ UK/ NL]
genotyping complete
statistical analyses currently underway

Concept

**Genetic Susceptibility to Obesity after
Childhood ALL**

PI

Kala Kamdar, Baylor

Genetic Susceptibility to Obesity after Childhood ALL

Aims

Using GWAS, determine relation between treatment-related obesity and genetic polymorphisms (SNPS and CNVs) after ALL

Replicate top SNPs and CNVs from discovery phase in an independent sample of ALL survivors from TCH

Funding

Leukemia Lymphoma Society

Status

GWAS analysis complete (as part of NCI GWAS initiative)

Statistical analysis currently underway

Replication cohorts being identified

Epigenomic Profiling of Metabolic Outcomes

Philip Lupo (TCCC) – funding: CCSS CDA

Specific Aim 1: Determine if gene-specific DNA methylation status is associated with obesity in ALL survivors by conducting genome-wide DNA methylation profiling

Methodology *DNA methylation profiles using a genome-wide approach among 48 obese ALL survivors and 48 normal-weight ALL survivors*

Specific Aim 2: Identify obesity susceptibility genes in ALL survivors through an integrated genomic and epigenomic analysis

Methodology *Perform methylation quantitative trait loci (mQTL) analyses to explore genotype-epigenotype associations*

Epigenomic Profiling of Metabolic Outcomes

Philip Lupo (TCCC) – funding: CCSS CDA

Status Validate findings of differential methylated loci (DMLs)

- **Conduct an integrative genomic analysis of obesity in ALL survivors**
 - Utilize existing methylation data and genome-wide SNP array data from CCSS to identify variants associated with DMLs.
 - Test the association between those SNPs and obesity among ALL survivors.
 - Test hypothesis that alleles associated with variable methylation will be enriched among obese ALL survivors
- **Determine if DMLs influence gene expression**
 - Evaluate the role of DMLs on gene expression in samples from an external cohort.
- **Develop an integrative genomic obesity risk prediction model**
 - Leverage CCSS to develop and validate a risk prediction model for obesity in those treated for ALL including clinical, genetic, epigenetic, gene expression data.

Concept

Neurofibromatosis and subsequent malignancies

PI

Smita Bhatia, UAB

Funding: P50-CA196519-01 (Clapp/ Shannon)

Neurofibromatosis and subsequent malignancies

Aim 1 Describe the magnitude of risk of SMNs in individuals with NF1.

Resources: CCSS and Tumor and NF1 Registries at CHOP

- i) children with NF1 and primary neoplasia (**NF1⁺cohort: CCSS and CHOP**);*
- ii) children with neoplasia, but without NF1 (**non-NF1cohort: CCSS**)*

Aim 1.1 Compare SMN risk in the **NF1⁺cohort** with SMN risk in **non-NF1 cohort**

Aim 1.2 Compare SMN risk in **NF1⁺cohort** exposed and not exposed to radiation and/or chemotherapy

Aim 2 Identify genetic alterations associated with radiation-induced tumorigenesis in individuals with NF1

Targeted exome sequencing of paired SMN/ germline tissue from individuals with NF1 to identify tumor-promoting pathways; characterize mutational landscape

Aim 3 Validate biologic importance of candidate pathways in radiation-induced tumorigenesis; determine whether radiotherapy promotes the development of MPNSTs.

mouse cell lines; mouse models

Status (score: 12) NOGA pending – will start study in Fall, 2015

Concept

**Genome-wide Association Study of
Subsequent Malignant Neoplasms among
Childhood Cancer Survivors (NCI/ CCSS)**

Genome-wide Association Study (GWAS)

- GWAS investigates the entire genome for common genetic variants in different individuals to see if any variant is associated with a trait
- GWASs typically focus on associations between single-nucleotide polymorphisms (SNPs) and traits like major diseases
- These studies compare DNA of two groups of participants: people with the disease (cases) and similar people without (controls).
- Each person gives a sample of DNA, from which millions of genetic variants are read using SNP arrays
- If one type of the variant (one allele) is more frequent in people with disease, the SNP is said to be "associated" with the disease.
 - The associated SNPs are then considered to mark a region of the human genome which influences the risk of disease.
 - GWASs cannot on their own specify which genes are causal.

Exome sequencing

- The **exome** is the part of the genome formed by exons.
 - Exons are sequences which when transcribed remain within the mature RNA after introns are removed by RNA splicing
- The exome of the human genome consists of ~180,000 exons constituting about 1-2% of the total genome, or about 30 megabases of DNA
- Though comprising a very small fraction of the genome, mutations in the exome are thought to harbor 85% of mutations that have a large effect on disease

Exome sequencing

- **Exome sequencing** is a technique for sequencing all the protein-coding genes in the genome (aka exome)
- It consists of first selecting only the subset of DNA that encodes proteins (aka exons), and then sequencing that DNA using a high throughput DNA sequencing technology
- The goal of this approach is to identify genetic variation that is responsible for both Mendelian and common diseases without the high costs associated with whole-genome sequencing

Whole Genome Sequencing

- **Whole genome sequencing** is a laboratory process that determines the complete DNA sequence of the genome at a single time
 - Using high-throughput genome sequencing technologies
- This entails sequencing all of the individual's chromosomal DNA as well as DNA contained in the mitochondria
- The tool of gene sequencing at SNP level is also used to pinpoint functional variants from association studies

GWAS vs. Whole Exome Sequencing

- GWAS is designed to identify common variants, contributing to common disease
 - These common variants are usually located in intronic regions of chromosome, which don't code for specific functional proteins
 - The associated SNPs are considered to mark a region of the human genome which influences risk of disease
 - GWASs cannot on their own specify which genes are causal
- To claim an association between gene and disease, these SNPs either have to lie hundreds of kilobases up- or downstream from a known causative gene, or the SNPs have to be in linkage disequilibrium with other SNPs of potential interest.
- A SNP's involvement is inferred due to its being in a place or associated with a gene known to have an effect on the disease of interest, much as a person found in a given place and hanging out with the right (or "wrong") crowd may become a "person of interest" in a criminal investigation.
- Whether the SNP is involved at all, or is a victim of "guilt by association," then becomes an issue that requires considerable probing.

GWAS vs. Whole Exome Sequencing

- By contrast, whole-exome sequencing looks at exonic regions of the chromosome, which code for functional proteins.
- It is no longer a matter of keeping bad company, but evidence of how the "crime" was committed, i.e., exactly how the protein was rendered nonfunctional and *why* a patient might be susceptible to or actually have the disease in question.
- WES is more likely to identify mutations -- often rare -- that have a greater impact on disease: ...80% to 90% of all [inherited] disease-causing mutations are located within protein coding regions
- So just by looking at 1% of the genome [using whole-exome sequencing]...one can identify almost 90% of disease-causing mutations.
- Using this logic, whole-exome sequencing would seem to be a better way to identify true disease-gene associations, and result in a savings of both time and money.

Concept

**Genome-wide Association Study of
Subsequent Malignant Neoplasms among
Childhood Cancer Survivors (NCI/ CCSS)**

GWAS resource for Genetic Investigation

- **Collaboration with Division of Cancer Epidemiology and Genetics**
 - **Genotyping platform**
 - Illumina® HumanOmni5-Quad BeadChip (~4.1 million SNPs)
- **Number of samples genotyped**
 - 5324 (European descent); 415 (non-European ancestry)
- **Aims**
 - Identify genetic variants that modify effect of RT and chemo on risk of subsequent neoplasms, and of risk independent of treatment exposure
- **Analyses**
 - Replication ongoing for promising SNPs

Concept

Exome sequencing to discover genetic variants that predispose childhood cancer survivors to the development of subsequent neoplasms (NCI/CCSS)

Genetics Working Group

Current Priorities

Vision for next 5 years

Genetics Working Group

Current priorities

1. Facilitate completion/ publication of studies currently underway
2. Replication of GWAS/SMN study

Vision for next 5 years

1. GWAS
 1. Facilitate approval of concepts in the queue pending primary GWAS analysis
 2. Encourage/ facilitate use of GWAS data for non-SMN outcomes
2. Whole Exome sequencing
3. Somatic mutational landscape/ pathogenesis of SMNs

Request for Proposals

use of GWAS data and corresponding outcomes-related data

- CCSS leadership is announcing a request for proposals to collaborate with CCSS and NCI investigators in the **use of GWAS data and corresponding outcomes-related data** to address innovative research questions relating to potential genetic contributions to risk for treatment-related outcomes.
- Requests will be accepted on an ongoing basis
- Requests will be reviewed three times per year, in February, June and October
 - cut-off for receipt of applications Feb1st, June 1st, Oct 1st

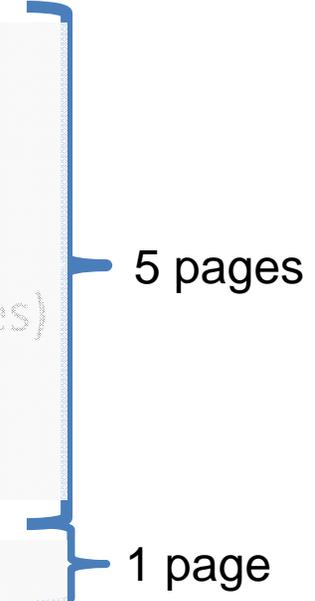
Request for Proposals

- All GWAS data will be posted on dbGap with following covariates
 - Age at diagnosis
 - Sex
 - Race/ ethnicity
- Researchers desiring additional covariates or linkage of genotypes to other phenotypes must submit a full application.
 - All projects will be viewed as a collaboration with one or more investigators from the CCSS and NCI being part of the research team.
 - All linkage of genotype and phenotype data (beyond that posted on dbGap) will only be carried out by the CCSS Data and Statistics Center.
- Data analyses may be performed by qualified individuals proposing the research, by CCSS biostatisticians, or both
 - All analyses and results must be vetted by CCSS biostatisticians prior to submission for publication.

Request for Proposals

- Investigators proposing to use GWAS data will comply with the following (by signing Data Use Certification)
 - Limiting use of data to the project described in Application for Data Access;
 - Not distributing data beyond those directly engaged in the analysis outlined in the Application for Data Access
 - Not attempting to identify or contact study participants from whom phenotype/genotype data were collected
 - Adhering to policies on timeframe for publications stemming from data
 - Providing annual detailed reports
 - Institutional documentation that CCSS material transfer agreement is acceptable
- Any violation of these policies can result in withdrawal of the project with return of CCSS data

Request for Proposals

- PI proposing to use GWAS data must submit application for Data Access to Greg Armstrong
 - Specific Aims
 - Background/significance of research question
 - Requested data elements
 - Approach (plans for validation, functional studies)
 - Statistical considerations
 - Timeline
 - Cited Literature
 - Investigative team and qualifications
 - NIH Biosketch for each investigator
 - Available resources/funding plans
- 
- 5 pages
- 1 page

Request for Proposals – Review Process

- **Applications will undergo a two-stage review process**
 - The initial review will be conducted by the CCSS Genetics Working Group Data Access Committee
 - Those applications approved by the Genetics Working Group Data Access Committee will be submitted to the CCSS Research and Publications Committee for review and comment

Request for Proposals – Review

- **Review Criteria**
 - Significance/impact of the research question
 - Approach (methodological, technical and statistical)
 - Innovation
 - Overlap with previous or current CCSS research
 - Qualifications of the investigative team
 - Concordance with CCSS priorities
 - Compliance with CCSS policies
 - Funding/ available resources
 - Timeline for manuscript submission
- **Reviewed applications will be categorized as**
 - approved
 - approved with stipulations
 - not approved, but a revised application may be reconsidered
 - disapproved, no future consideration.

Request for Proposals – Post Approval Process

- The PI and his/her research team will work together with CCSS collaborators assigned to the project to facilitate construction of the analytic data set, establish a task-specific timeline, agree upon a tentative list of target journals, and authorship.
- A written progress report will be required annually until the project has been completed (i.e. publication).
 - CCSS reserves the right to withdraw approval for any project that fails to demonstrate adequate progress (i.e., if the project is delinquent by >12 months after the designated timeline).
- Review and approval of results by the CCSS Genetics Working Group Chair and the assigned CCSS biostatistician is required prior to any presentation or publication of the data from the project.

Request for Proposals

Data Access Committee

- Smita Bhatia (chair)
- Greg Armstrong
- Les Robison
- Yutaka Yasui
- Lindsay Morton
- Joshua Sampson
- Wendy Leisenring
- Stephen Chanock (ad hoc)
- Peggy Tucker (ad hoc)
- Working Group Chairs (ad hoc)
- NCI DCEG investigators (ad hoc)

Genetics Working Group

Concepts in the queue pending
primary GWAS analysis

Concept

Radiation-related thyroid cancer

PI

Yuri Nikiforov, U of Pittsburgh

Radiation-related thyroid cancer

- Aim** Test whether alterations in DNA repair genes (*ATM*, *BLM*, *NBS1*, *DNA-PKcs*, *Ku70*, *XRCC4*, *RAD51*) lead to thyroid cancer after radiation
- Design** Matched case-control study design
Matching criteria: Primary diagnosis, radiation field to involve thyroid, sex, age at exposure, race/ ethnicity, duration of follow-up to exceed latency between primary diagnosis and thyroid cancer
- Status** Cases and controls identified
Samples to be released post-NCI GWAS effort

Genetic Susceptibility to SMNs

Onel/ U of Chicago

Specific Aim	Test association of top SNPs from GWAS study with risk for second cancers after primary diagnoses other than HL
Eligibility criteria	All patients with SMNs after primary cancer other than HL, and controls 1:1 matched for primary cancer, latency, age of treatment for primary cancer, treatment modality, gender, and race/ethnicity
Status	Request awaiting completion of primary analysis of NCI/CCSS GWAS

PRDM1 and somatic mutations in SMNs after HL

Onel/ U of Chicago

Specific Aim	Genotype chr 6q21 risk locus in SMN tissue samples <i>Assess (FISH/ IHC) PRDM1, MYC mutation status in SMNs</i>
Eligibility criteria	Archival tumor samples of radiation-induced SMNs after HL
Status	Concept is being finalized

Concept

Genetic Susceptibility to anthracycline-related CHF – Replication study

PI

Smita Bhatia, UAB

Genetic Susceptibility to anthracycline-related CHF

Replication study

- Aims** Replicate significant findings identified in the Discovery set (using COG-case-control study [ALTE03N1]) in an independent case-control set from CCSS
- Status** Concept approved
Cases and controls identified for validation
Samples to be released post NCI GWAS effort

Where do we want to go?

- The need of the hour...

....Precision Medicine

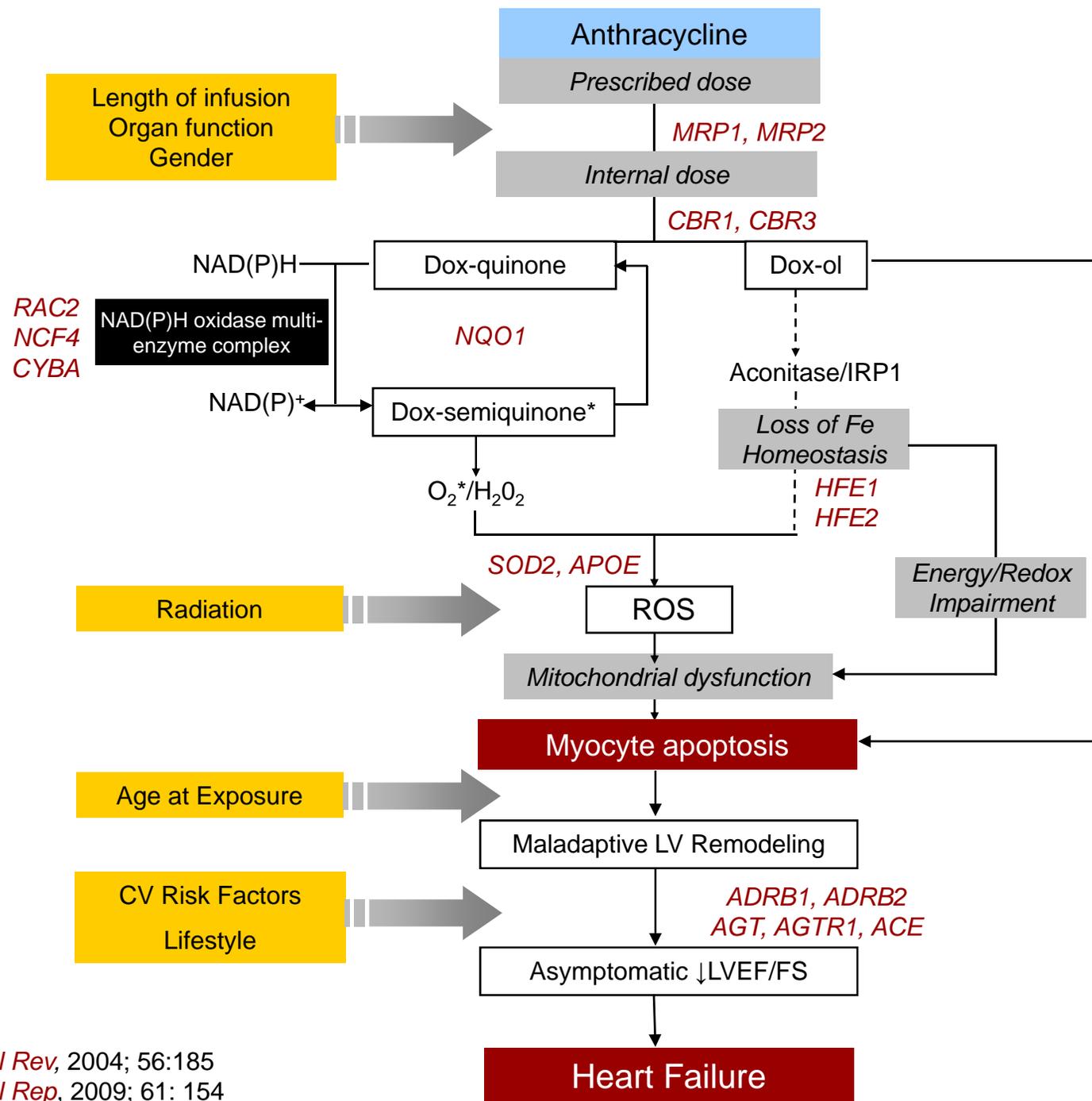
DNA repair in multiple SNs

Bhatia, UAB

Specific Aim Examine role of candidate genes involved in multiple SNs

Status AOI approved

Concept awaiting primary analysis of NCI/GWAS study



Modified:

Pharmacol Rev, 2004; 56:185

Pharmacol Rep, 2009; 61: 154

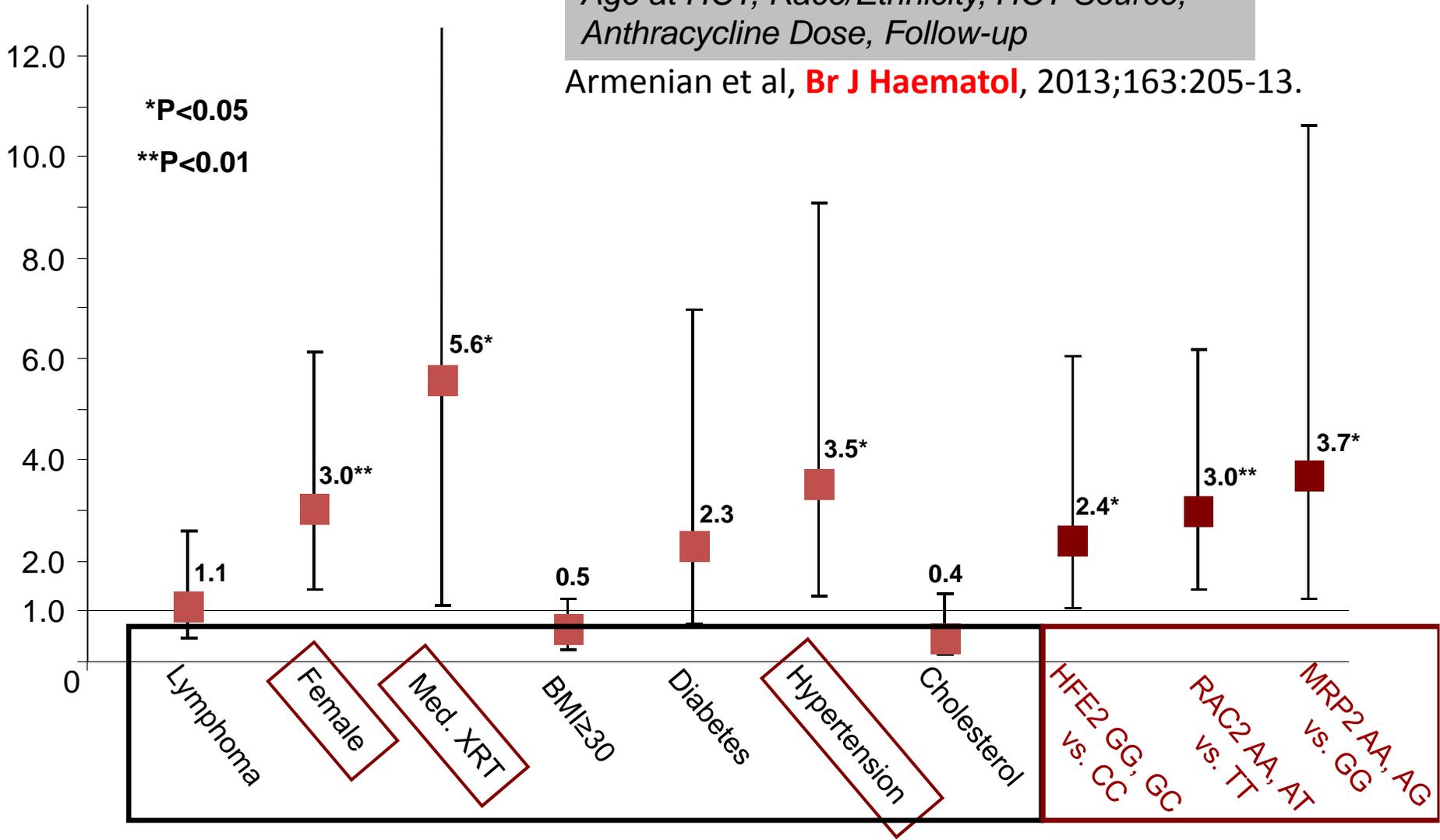
Risk of CHF

Case-Control Matching:

Age at HCT, Race/Ethnicity, HCT Source,
Anthracycline Dose, Follow-up

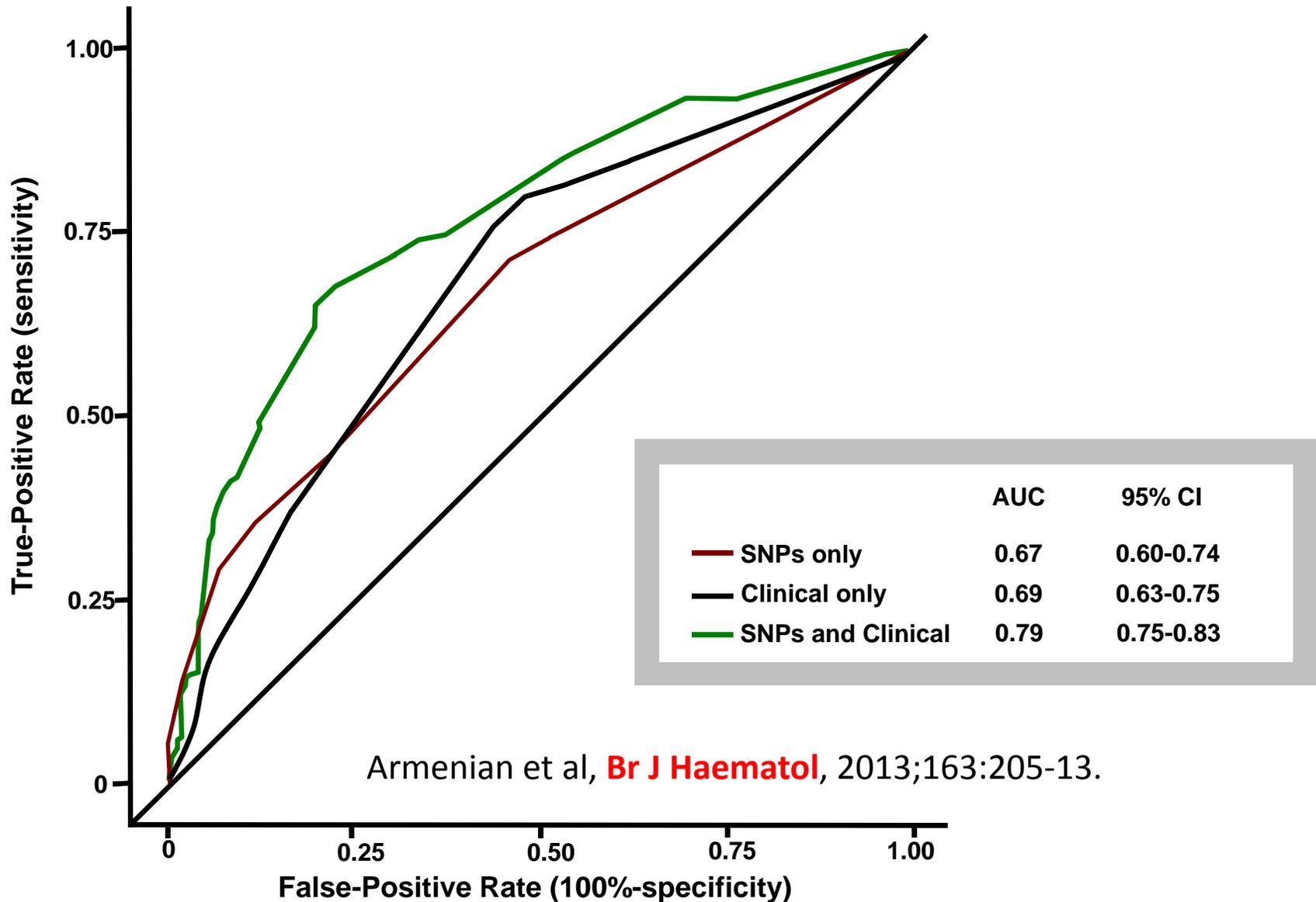
Armenian et al, **Br J Haematol**, 2013;163:205-13.

Odds Ratio



*P<0.05
**P<0.01

Receiver operating curve (ROC)



Questions?