Translational Bioinformatics for Immunogenomics

Table of contents

Welcome	5
Introduction	6
Human Leukocyte Antigens	7
Background	7
HLA Class I	7
HLA Class II	7
HLA Nomenclature	7
Functional Divergence	7
HLA Imputation Programs	8
Killer Cell Immunoglobulin-like Receptors	9
Background	9
KIR Locus	9
KIR Diversity	12
NK Cell Education	13
KIR Nomenclature	13
Inhibitory KIR	13
Activating KIR	14
Broad KIR Haplotypes	14
KIR Ligand Motifs	14
KIR3DL1 and KIR3DS1	15
KIR Allele Imputation Programs	16
ERAP	17
Epistatic Interactions	18
KIR-HLA	18
Bonus	19

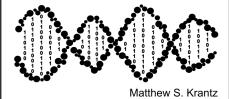
Drug Alloway	0
6 6	20
0 0,	
9	20
<u> </u>	0.0
0	20
v 0 0v	20
8	20
<u>o</u>	21
0	22
1 0	23
	23
1 1 0	23
Fluoroquinolones 2	23
Macrolides	23
Penicillins	25
Sulfa Antibiotics 2	25
Vancomycin	25
	25
	25
	25
9	26
	26
	27
0	27
8	27
	28
	28
	28
<u>o</u>	28
	28
v vi	29
0	29
3	29
Management	29
Genotype Imputation 3	0
	80
-	3 0
*	0 80
	81
	11

Bioinformatic Best Practices	32
Project Organization	32
Version Control with Git	32
File Naming Conventions	33
Application Containers with Docker	34
Presenting Your Medical Research	35
Font	35
Font Size	36
Word Count	36
Timing	36
Figures	36
References	37
Equipment	37
Laptop	37
Hub	37
USB Drive	38
Presentation Remote	38
On Being a Physician-Scientist	39
Daily Writing Practice	39
Research on Daily Writing	40
Research on Daily Writing	40 40
Time Target for Daily Writing	40
Time Target for Daily Writing	40 40
Time Target for Daily Writing	40 40 41
Time Target for Daily Writing	40 40 41 41
Time Target for Daily Writing	40 40 41 41 41
Time Target for Daily Writing Tips for Daily Writing	40 40 41 41 41 42
Time Target for Daily Writing Tips for Daily Writing Benefits of Daily Writing Academic Medicine Jobs AAMC Faculty Salary Report Tenure-Track Offer Letters NIH Loan Repayment Program	40 40 41 41 41 42 42
Time Target for Daily Writing Tips for Daily Writing Benefits of Daily Writing Academic Medicine Jobs AAMC Faculty Salary Report Tenure-Track Offer Letters NIH Loan Repayment Program Online Resources	40 40 41 41 41 42 42 42
Time Target for Daily Writing Tips for Daily Writing Benefits of Daily Writing Academic Medicine Jobs AAMC Faculty Salary Report Tenure-Track Offer Letters NIH Loan Repayment Program Online Resources Edge for Scholars	40 40 41 41 41 42 42 42 42
Time Target for Daily Writing Tips for Daily Writing Benefits of Daily Writing Academic Medicine Jobs AAMC Faculty Salary Report Tenure-Track Offer Letters NIH Loan Repayment Program Online Resources Edge for Scholars NCFDD	40 40 41 41 41 42 42 42 42 42
Time Target for Daily Writing Tips for Daily Writing Benefits of Daily Writing Academic Medicine Jobs AAMC Faculty Salary Report Tenure-Track Offer Letters NIH Loan Repayment Program Online Resources Edge for Scholars NCFDD Professional Organizations Suggested Readings Not Discussed	40 40 41 41 42 42 42 42 42 42
Time Target for Daily Writing Tips for Daily Writing Benefits of Daily Writing Academic Medicine Jobs AAMC Faculty Salary Report Tenure-Track Offer Letters NIH Loan Repayment Program Online Resources Edge for Scholars NCFDD Professional Organizations Suggested Readings	40 40 41 41 42 42 42 42 42 42 42

Welcome

This is the website for "Translational Bioninformatics for Immunogenomics."

- > Translational
- > Bioinformatics
- > for
- > Immunogenomics



Introduction

Human Leukocyte Antigens

Background

HLA is located on chromosome 6 in the Major Histocompatibility Complex (MHC).

HLA Class I

HLA class I molecules are expressed by healthy nucleated cells.

HLA Class II

HLA class II molecules are expressed by professional antigenpresenting cells (APC)-dendritic cells, macrophages, and B cells.

HLA Nomenclature

Functional Divergence

Heterozygosity of HLA class I genes is associated with better outcomes after HIV infection. This is thought to be due to a greater repertoire of HIV peptides presented and cytotoxic T cell response. However, looking at HLA class I allotype alone does not take into account differences in actual peptide repertoire. Viard et al. (2024) developed a metric to measure this difference, termed "functional divergence." Functional divergence predicts the peptide repertoire as a continuum. They

showed that greater functional divergence was associated with better HIV outcomes. Functional divergence may be relevant to other diseases where HLA heterozygosity confers advantage, such as infection, vaccination, and immunotherapy.

You can download functional divergence estimates for pairwise combinations of HLA-A, HLA-B, and HLA-C alleles from their article's Supplementary Materials. The functional divergence measure ranges from 0 (i.e., smallest functional divergence) to 1 (i.e., greatest functional divergence).

HLA Imputation Programs

	Programming			
	Lan-	Input		
Name	guage	Data	Output	Reference
SNP2HLA	Commai	ndPLINK	HLA	Jia et
	line in-	binary	class I	al.
	terface	for-	and II	(2013)
		mat	alleles	
HIBAG	R	Plink	HLA	Zheng
		binary	class I	et al.
		for-	and II	(2014)
		mat	alleles	

Killer Cell Immunoglobulin-like Receptors

Background

KIR is located on chromosome 19 (19q13.4) in the Leukocyte Receptor Complex (LRC). KIR is expressed on the surface of Natural Killer (NK) cells and some T cells. KIR do not undergo somatic rearrangement—a key difference from T-cell receptors. KIR interacts with HLA class I—their cognate ligand—to recognize and destroy unhealthy tissue cells while preventing the same from occurring to healthy cells. Therefore, NK cells play a role in fighting infections, resisting some cancers, pregnancy, and preventing autoimmunity. For further reading and references, I highly recommend the review article by Pollock, Harrison, and Norman on the immunogenetics and co-evolution of KIR and HLA class I.

KIR Locus

Adapted from Pollock, Harrison, and Norman. JACI: In Practice. 2022.

Gene

3DL3

2DS2

2DL2/3

2DL5B

2DS3

2DL1

2DL4

3DL1

3DS1

2DL5A

2DS5

2DS4

2DS1

3DL2

Function

Inhibit.

Activ.

Inhibit.

Inhibit.

Activ.

Inhibit.

Activ.

Inhibit.

Activ.

Inhibit.

Activ.

Activ.

Activ.

Inhibit.

Alleles

228

65

98

Allotypes

HLA Class I Ligand Motifs

```
B7H7
A*11 C1
B46:01 B73:01 C1 C2
PVR
?
C2
HLA-G
Bw4+ HLA-A and Bw4+ HLA-B
Bw4+ HLA-B and HLA-F
PVR
C2
A*11 HLA-C
C2
A3 A11
```

KIR Diversity

Practice. 2022.

KIR diversity is influenced by gene content variation and sequence variation. Distinct DNA sequences of KIR genes are called "alleles." Distinct polypeptide sequences of KIR genes are called "allotypes." Because different DNA sequences of KIR gene can lead to the same polypeptide, there are more alleles than allotypes for a given KIR gene.

: Adapted from Pollock, Harrison, and Norman. JACI: In

KIR Di-	
versity Con-	
cept	Definition
Gene	Presence/absence, fusion, duplication
Content	
Varia-	
tion	
Sequence	May alter ligand affinity or specificity, signal
Varia-	transduction ability, or surface expression (e.g.,
tion	promoter activity, translation, intracellular
	trafficking)
Allele	Distinct DNA sequence
Allotype	Distinct polypeptide sequence

NK Cell Education

NK Cell		
Education		
(i.e.,	Corresponding Pairs of	Cytotoxicity and
Arming,	KIR and HLA Class I	other Effector
Licensing)	Ligands	Abilities
Strong	Many	More
Weak	Few	Less

KIR Nomenclature

Inhibitory KIR

The main role of inhibitory KIR is to prevent cytotoxic NK and T cells from killing tissue cells—unless their HLA class I expression is lost or altered by infection or mutagenesis.

Activating KIR

Activating KIR help identify diseased cells for destruction by cytoxic NK and T cells. Binding of foreign peptides by HLA class I molecules retained by infected cells may be most critical for activating KIR.

Broad KIR Haplotypes

Broad KIR Haplotype	KIR Copy Number Variation	KIR Gene Organization	Activating KIR
A	Relatively stable	Generally non-variable	Less
В	Extensive	Highly variable	More

KIR Ligand Motifs

Table 0.4: Adapted from *Pollock, Harrison, and Norman.*JACI: In Practice. 2022.

KIR			
Lig-			
and			
Mo-	HLA-A		HLA-C
tif	Allotypes	HLA-B Allotypes	Allotypes
$\overline{A3/A}$	11A*03, A*11		

KIR			
Lig-			
and			
Mo-	HLA-A		HLA-C
tif	Allotypes	HLA-B Allotypes	Allotypes
Bw4	A*23, A*24,	B*07:27, B*08:02,	
	A*32	B*08:03, (B13), B*15:13,	
		B*15:16, B*15:17,	
		B*15:23, B*15:24,	
		B*15:36, B*15:43,	
		B*15:67, B*27:01,	
		B*27:02, B*27:03,	
		B*27:04, B*27:05,	
		B*27:07, B*37, B*38,	
		B*40:13, B*40:19, B*44,	
		B*47, B*49, B*51, B*52,	
		B*53, B*56:07, B*57,	
		B*58, B*59	
C1	C*01, C*03,	B*46, B*73	
	C*07, C*08,		
	C*12:02,		
	C*12:03,		
	C*12:06,		
	C*12:08, C*13,		
	C*14, C*16		
C2	C*02, C*03:07,		
	C*04, C*05,		
	C*06, C*12:04,		
	C*12:05,		
	C*12:07,		
	C*14:04, C*15,		
	C*16:02, C*17,		
	C*18		

KIR3DL1 and KIR3DS1

Because of significant non-allelic recombination in the KIR region, the distinction between KIR genes and alleles can be confusing. Specifically, KIR3DL1 and KIR3DS1 are alleles of the

same gene. Of the KIR3DS1 allotypes–3DS1 θ 13 and 014–are observed with the greatest frequency in any population.

KIR Allele Imputation Programs

	Progr	amm	ing	
	Lan-	Inpu	t	
Name	guage	Data	Output	Reference
PONG	R	PLIN bi-	KKIR3DL1/S1 alleles (Global Model	Harrison, 2022
		nary	includes 51 alleles)	
		for-		
		$_{\mathrm{mat}}$		
KIR*IMP	Online	e HAP	S/19/AMIR Itypes: 17 loci	Vukcevic,
	por-	for-	(presence/absence	2015
	tal	$_{\mathrm{mat}}$	and copy number)	
			plus 2 extended	
			haplotype	
			classifications (A	
			and B haplotypes)	

ERAP

ERAP is located on chromsome 5.

Epistatic Interactions

KIR-HLA

Epistatic interactions between KIR and HLA are associated with ankylosing spondylitis (Hanson, 2020)

Bonus

Drug Allergy

Immediate Drug Allergy

Skin Testing

Skin Prick Testing

Intradermal Testing

Delayed Drug Allergy

Skin Testing

Table 0.1: Utility of patch and intradermal skin testing for delayed drug allergy reaction types

Patch Testing	Intradermal Testing
Useful if positive	Useful if positive
Useful if positive	Useful if positive
Oberui ii positive	Obeital if positive
Low sensitivity	Contraindicated
but potentially	due to concern for
useful if positive	potential
	reactivation
Useful if positive	Useful if positive
	Useful if positive Useful if positive Low sensitivity but potentially useful if positive

Reaction	Patch Testing	Intradermal Testing
Fixed drug eruption	Useful if applied to the site of reaction	Not useful
Drug-induced organ injury (e.g., kidney, liver)	Not useful	Not useful

! Important

No delayed skin testing method has 100% negative predictive value.

Table 0.2: Shared characteristics of patch and intradermal testing $\frac{1}{2}$

Characte	Characteri Dict ails					
Timing	Perform at least 6 to 8 weeks after reaction; and 6 months or later after DRESS					
medica-	almost medications okay to continue, including anti-histamines and beta-blockers. Should be off of					
tions	steroids for 1 month or prednisone equivalent dose 10 mg/day					

Intradermal Testing

Table 0.3: Characteristics of intradermal testing

Characteristic Details				
Testing site	Volar forearm or extensor upper arm			
Testing	Must be sterile; often higher concentrations			
reagents	than those used for immediate skin testing			
Reading	At 24 hours			
Controls	+ None			
	- Saline			

Characteristic Details				
Test interpretation	+ Papule present - Negative			

Patch Testing

Table 0.4: Characteristics of patch testing

Characteristic	Details			
Testing site	Back or upper arm (needs to be hairless)			
Testing	1% and $10%$ of reagent grade product; $10%$			
reagents	and 30% of trade product; most commonly			
	used vehicle is petrolatum			
Controls	+ None			
	- Petrolatum			
Shelf-life of	Most antibiotics at room temperature are			
patch test	stable for 1 to 3 months; check with USP			
mixes	Pharmacopeia for verification			
Patches	Finn chambers (can be aluminum or molded			
	plastic)			
Tape	Use hypoallergenic paper tape			
Reading	At 48 hours (85% of drugs-if will be			
	positive-are positive by this point); 72 hours;			
	96 hours; and 1 week			
Test interpre-	- Negative			
tation	? Doubtful reaction			
	+ Weak reaction, erythema			
	++ Strong reaction, erythema, papules, or			
	vesicles			
	+++ Extreme, bullous, ulcerative			

Specific Drugs

Antibiotics

Cephalosporings

Fluoroquinolones

Macrolides

Background

What is the chemical structure of macrolides?

Macrolides are defined by a large lactone ring, which varies from 12 to 16 carbons, with 1 or more attached sugar chains. Eyrthromycin and clarithromycin have 14 carbons in their lactone rings while azithromycin has 15.

What is the mechanism of action of macrolides?

As 50S ribosomal subunit inhibitors, macrolides exert their bacteriostatic effect by inhibiting protein synthesis.

What is the cross-reactivity pattern macrolide antibiotics?

While not extensively studied, macrolide antibiotics with a different number of carbon atoms in their lactone ring are tolerated by most patients. Macrolide antibiotics are also unlikely to cross-react with macrolide immunosuppressants (e.g., tacrolimus, sirolimus).

What are some infections that use macrolides as first-line therapy?

Clarithyromycin is used as part of combination treatment for *H. pylori*. Azithromycin is a part of the first-line combination therapy for *Mycobacterium avium* complex.

Non-Hypersensitivity Reactions

GI Side Effects

Because macrolides are also agonists for the motilin receptor—stimulating gastric and small intestine motility—they can cause nausea, vomiting, diarrhea, and abdmoninal cramping. Accordingly, erythromycin can be used as a treatment for gastroparesis.

Sensorineural Otoxicity

Macrolides can cause usually, transient sensorineural ototoxicity.

QT Interval Prolongation

Macrolides are associated with QT interval prolongation.

Skin Testing

Skin testing for macrolide immediate hypersensitivity has not been shown to be reliable.

Oral Challenge

For patients with a history of immediate hypersensitivity reaction to a macrolide, graded azithromycin challenge can be performed, starting with azithromycin 25 mg followed by 1 hour observation then 250 mg followed by 2 hour observation.

For patients with a history of non-severe delayed hypersensitivity reaction to a macrolide, single dose azithromycin 250 mg challenge with 2 hour observation can be performed. Patients should be instructed to report any other delayed symptoms, which may occur up to 24 to 48 hours after the challenge dose.

Penicillins

Sulfa Antibiotics

Vancomycin

Buproprion

Iron

Background

Why is intravenous (IV) iron used?

IV iron is used for the treatment of iron deficiency anemia when oral iron is effective or not tolerated.

What IV iron formulations are available in the United States?

Formulations available in the United States include low-molecular-weight iron dextran (LWMID), ferric gluconate, iron sucrose, ferumoxytol, iron isomaltoside, and ferric carboxymaltose.

Note

High-molecular weight iron dextrans were discontinued in the United States due to having a higher rate of hypersensitivity reactions.

Table 0.5: Characteristics of iron formulations available in the United States

	Iron	Iron		Ferric	Iron	
Generic	glu-	Su-	ca	arboxy-	isomal-	
name	conate	crose	LMWh	Paltose	toside	Ferumoxytol
Brand	Ferrlecit	Venofer	:INFe D	jectafer	Monofer	FeraHeme
name						
Molecular	289-	30-60	165	150	150	750
weight (kD)	440					

	Iron	Iron		Ferric	Iron	
Generic	glu-	Su-		carboxy-	isomal-	
name	conate	crose	LMWhaltose		toside	Ferumoxytol
Labile iron (% injected dose)	3.3	3.5	2	0.6	1	0.8

Immediate Hypersensitivity Reactions

What is the incidence of anaphylactic reactions with IV iron?

Anaphylactic reactions—when high-molecular weight dextrans are excluded—occur with an incidence of < 1 in 200,000.

Significant differences in reaction risk have not been shown among low-molecular weight iron dextran, iron sucrose, ferric gluconate, and ferric carboxymaltose.

What is mechanism of most IV iron immediate hypersensitivity reactions?

Most IV iron immediate hypersensitivity reactions are mediated through complement-activation related pseudoallergy (CARPA). Rarely, hypersensitivity reactions are IgE-mediated.

Minor Infusion Reactions

What are the symptoms of minor infusion reactions to IV iron?

Symptoms of minor infusion reactions to IV iron include—flushing, chest/back tightness, myalgias—and, importantly, do not have any features of anaphylaxis.

What is considered to be the main driver of minor infusion reactions to IV iron?

Labile, or also called "free," iron is associated with minor infusion reactions to IV iron.

Skin Testing

What is the utility of immediate skin testing for IV iron hypersensitivity reactions?

As most hypersensitivity reactions are non-IgE-mediated—rather via CARPA—skin testing has limited utility for evaluating IV iron hypersensitivity reactions; however, it may help detect the rare patients with IgE-mediated hypersensitivity.

Management

What are some approaches to subsequent IV iron administration in patients with previous IV iron reactions?

Approaches for patients with history of mild to moderate IV iron reactions include: switching to an alternative IV iron formulation, slowing the infusion rate (e.g., 10% of recommended rate during the first 10 to 15 minutes), and/or pre-treatment with non-sedating, second generation antihistamines.

For patients with a history of anaphylactic reactions to IV iron, desensitization can be considered, such as ferric gluconate.

References

Gómez-Ramírez S, Shander A, Spahn DR, et al. Prevention and management of acute reactions to intravenous iron in surgical patients. *Blood Transfusion*. Published online April 10, 2019. doi:10.2450/2018.0156-18

Muñoz M, Gómez-Ramírez S, Bhandari S. The safety of available treatment options for iron-deficiency anemia. Expert Opin Drug Saf 2018; 17: 149-59.

Local Anesthetics

Radiocontrast

Background

How do modern, radiocontrasts differ from older radiocontrasts?

Modern radiocontrasts are non-ionic, iodinated and either iso-osmolal or low-osmolality. Older radiocontrasts were high-osmolality—which are no longer used intravenously.

Is radiocontrast allergy related to iodine or seafood allergy?

The radiocontrast molecular structure is responsible for hypersensitivity reactions—not iodine or seafood allergy. Shellfish allergy is secondary to tropomyosin not iodine.

Infusion Reactions

These are also referred to as "toxic" or "chemotoxic" reactions. Characteristic symptoms include transient warmth/flushing, nausea/vomiting, chest pain, metallic taste, hypertension, and/or vasogal signs.

Immediate Hypersensitivity Reactions

The most common immediate hypersensitivity reaction to radio contrast is mild urticaria and pruritus, occuring in ${\sim}0.9\%$ -3.1% of patients receiving radio contrast. Anaphylaxis occurs in 0.02% - 0.04% of patients. Of immediate hypersensitivity reactions, 70% occur within 5 minutes of radio contrast injection and 96% of severe reactions occur within 20 minutes.

Delayed Hypersensitivity Reactions

The most common delayed hypersensitivity reaction to radiocontrast is a maculopapular exanthem–occurring in 1 to 3% of patients.

Skin Testing

Skin prick testing to the culprit and other radiocontrasts followed by intradermal testing—if skin prick testing is negative—can be useful for identifying an alternative radiocontrast agent. A skin test negative radiocontrast alternative has a 95% NPV.

Intravenous Challenge

Intravenous challenge can be performed to radiocontrast with various protocols—such as 1 mL, 5 mL, 15, mL, and 50 mL (cumulative 71 mL) at 60 minute intervals.

Management

The culprit radiocontrast should be avoided, and an alternative radiocontrast should be used, guided by negative skin testing if available. Other measures to decrease risk of recurrent radiocontrast reaction include: lowering the radiocontrast dose, decreasing the injection speed, and pre-treatment with non-sedating, second generation antihistamines and/or corticosteroids.

Genotype Imputation

Michigan Imputation Server

The Michigan Imputation Server is a free next-generation genotype imputation platform. You can learn more about the Michigan Imputation Server by visiting their Getting Started documentation. The 1000 Genomes Phase 3 (Version 5) Reference Panel is available on the Michigan Imputation Server.

TOPMed Imputation Server

The TOPMed Imputation Server is another free next-generation genotype imputation platform developed by the University of Michigan and powered by data from the TOPMed Program investigators. You can learn more about the TOPMed Imputation Server by visiting their Getting Started documentation. The TOPMed Version 3 Reference Panel was released in December 2023.

Reference Panels

	Geno	me			
	As-	No. o	f Sites	S	
	sem-	Sam-	(chr]	L-	
Reference Panel	bly	ples	22)	Chr	: Imputation Server
1000 Genomes	GRC	h 3 75/042	g 14 99,1	431,60	5Michigan
Phase 3				22,	Imputation Server
(Version 5)				X	

	Genor As-	me No. o	f Sites	3	
	sem-	Sam-	(chr1		
Reference Panel	bly	ples	22)	Chr	:.Imputation Server
TOPMed	GRCl	1 38 /3 5/2	9378 45,0	60D <u>,</u> 1	84TOPMed
(Version 3)				22,	Imputation Server
				X	

Genome Assemblies

The Genome Reference Consortium (GRC) is the main source of human genome assembly data. The most recent human genome assembly version is GRCh38, released in 2013. The "h" in "GRCh" stands for "human." The GRC also maintains genome assembly data for rat (r), mouse (m), zebrafish (z), and chicken (g for gallus). Major updates, called "versions", are released every few years. Minor updates are called "patches" and are released more frequently.

GRCh38 is referred to as "hg38" in the University of California Santa Cruz (UCSC) Genome Browser. The "hg" stands for "human genome." Before the GRCh38 genome assembly, the version numbers of the GRC and UCSC Genome Browser genome assemblies did not match. For example, when the GRCh37 genome assembly was released in 2009, the UCSC Genome Browser version was "hg19." Therefore, to minimize confusion, starting with the GRCh38 genome assembly, the UCSC Genome Browser version number was matched as "hg38."

GRC Version	UCSC Version	Year Re- leased	Genome Coverage	Alternate Haplotypes
GRCh37	hg19	2009	~92.5%	3 regions with 9 alternate loci
GRCh38	hg38	2013	95%	178 regions with 261 alternate loci

Bioinformatic Best Practices

I recommend the tutorial, "A Reproducible Data Analysis Workflow With R Markdown, Git, Make, and Docker" as a starting point for R-based data analyses (Peikert & Brandmaier, 2021).

Project Organization

Bash Commands to Create Folder Directory Structure for Your R Project

```
cd </path/to/parent/directory>
mkdir <your-r-project-folder>
cd <your-r-project-folder>
touch README.md
mkdir data doc src bin outputs
```

Once you have downloaded your raw data to your data folder, you should make the contents of the data folder read-only (non-editable) with the following command: $chmod\ u-w\ -R$ data/

Version Control with Git

I recommend the Using Git and GitHub with RStudio Cheatsheet for additional helpful commands.

Verify Git Installation and Version

```
which git # request path to your Git executable
git --version # check your Git version
```

Introduce Yourself to Git

```
git config --global user.name "<username>"
git config --global user.email "<email>"
```

Create a New Repository on GitHub

Go to GitHub to create your new repository, then initialize your repository from the command line.

```
cd </path/to/your-r-project-folder>
echo "# your-r-project-folder" >> README.md
git init
git add README.md
git commit -m "first commit"
git branch -M main
git remote add origin https://github.com/<user.name>/<your-repository>.git
git push -u origin main
```

Add, Commit, and Push Files to Remote Repository

```
git add <file-name>
git commit -m "description"
git push
```

File Naming Conventions

In your README.md, you should define naming conventions for your project files. The main elements for a file naming convention are metadata, separator, and version tracking. I recommend the File Naming Conventions Worksheet (Briney, 2020) to develop your file naming conventions.

Metadata	Separator	Version Tracking
3 to 5 pieces max (e.g. sam- ple ID, date in ISO 8601 format such as YYYY- MM- DD)	Dashes (-), underscore (_), or camel case (i.e., capitalize each word without spaces)	Numeric (e.g., v01) or Status (e.g., raw, processed)

i Example

My naming convention for R Markdown analysis files is: "analysis-YYYY-MM-DD-version.Rmd" where version starts with "v01." This is my first analysis file, "analysis-YYYY-MM-DD-v01.Rmd" $^{\prime\prime}$

Application Containers with Docker

Presenting Your Medical Research

Font

You should use a sans-serif font like Arial to maximize readability. "Serifs" are extending features at the end of letters. Times New Roman is a serif font.

Table 0.1: Sans-serif versus serif fonts¹

Font	IllustratioExamples
Sansserif font	AaB bCc
Serif font	AaBbCc New Roman, Georgia,
Serifs (colored in red)	AaBbCc

Font Size

Slide Section	Font Size
Title	36 - 44
Text	24-28
(e.g., Bullets, Figures,	
Tables)	
References	20-24

Word Count

The fewer words, the better. A rule to follow is the 7×7 rule: no more than 7 lines and no more than 7 words per line.

Timing

You should estimate approximately 1 minute per slide.

Figures

I recommend creating your figures as Scalable Vector Graphics (SVG). The main advantages of the SVG format include always maintaining its resolution and smaller file size than pixel-based image formats (e.g., JPEG).

Some tools that you can use to get started creating SVG include Microsoft PowerPoint (subscription), Adobe Illustrator (subscription), draw.io (free), and Inkscape (free). Draw.io is best for diagrams and flowcharts. Inkscape is better for flexible drawings. Both draw.io and Inkscape are integrated with Bioicons, an open-source extension which includes >1700 icons for scientific illustrations.

¹Font images are recreated by User:Stannered, original by en:User:Chmod007 - en:Image:Serif and sans-serif 01.png, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=2058303

In Microsoft PowerPoint, you can create an SVG file by selecting all shapes, right-clicking, choosing "Save as Picture", and then picking "SVG" as the "Save as Type."

References

Cite references at the bottom of your slides as you present information.

i Format

Last Name. Journal Abbreviation. Year.

Equipment

Laptop

Bring your own laptop to presentations in case there isn't a desktop computer for you to use, or it is not functioning, reliable or frustratingly slow.

Hub

What is worse than not being able to connect your laptop to the correct cable? While a good host for a presentation should have a hub (or dongle if that's your preferred terminology), you can come prepared with your own too—particularly important if you have a laptop with only USB-C ports and no HDMI port.

There are lots of options for hubs. If you are looking for a recommendation, I've found that Anker usually has a selection of high-quality and affordable hubs.

USB Drive

Do you want the entire audience to see your most recent emails when you login to download the PowerPoint you emailed yourself? No. Me either. To avoid this, bring your presentation loaded onto a USB drive, which should ideally have both USB-A and USB-C ports. Or, you can also avoid this by using your own laptop—where the presentation should already be downloaded.

Presentation Remote

I don't feel as strongly about bringing your own presentation remote as your own laptop, hub, and USB drive—but I think it is another piece of equipment to consider. This helps keep you from being tethered to and white knuckling the podium during your talk.

One option to consider is the Logitech Spotlight Presentation Remote—which includes features such as magnification, vibration alerts for time management (e.g., 5 minutes left), 3 hours use from 1 minute of charging, and connection by USB receiver or Bluetooth—in addition to slide advancement.

On Being a Physician-Scientist

Daily Writing Practice

Why is it that the most important academic activity for tenure, promotion, and professional reputation—writing—has the least amount of built-in accountability?

If you are a physician-scientist, you are a writer; therefore, you should write everyday (Monday - Friday).

Table 0.1: Built-in accountability and importance for tenure, promotion, and professional reputation by activity

Less	Built-in accountability	More
•	->	•
Writing	Activity	Service
	<-	
• Articles		• Teaching
• Grants More	Importance for tenure, promotion, and professional reputation	Clinic/Consults $Less$

Note

The most important part of your promotion—writing—has the least accountability.

Table 0.2: Limiting beliefs to cultivating a daily writing practice. Adapted from NCFDD.

Belief	Reality
"I need huge blocks of time." "I must be inspired to write."	"Both unrealistic and untrue. You can productively write in 30 minute blocks!" "No, you don't. If you put it on your calendar, you can show up to write just like you show up to meetings you don't want to attend."
"Writing is what I do when I'm done thinking."	"Writing is thinking."

Research on Daily Writing

Table 0.3: Adapted from Boice et al

	Draft Pages Written
Participant Groups	per Year
No change	17
Wrote daily and recorded progress	64
Wrote daily, recorded progress, and	157
were accountable	

Time Target for Daily Writing

Your goal should be to spend 3 or more hours per week on scholarly writing. So, if you write 30 minutes Monday to Friday, you are already at 2.5 hours!

Tips for Daily Writing

Schedule your writing in your calendar like any other meeting or clinical duty.

You should write first thing in the morning. Knock out the most important daily task for your career first!

Map complex goals to attainable steps.

Use a timer, stop when the timer goes off (to avoid slipping back into writing in huge chunks).

Leave yourself a "breadcrumb," so you can pick up where you left off.

Give yourself a treat after writing.

Do a reflection at the end of the week (Friday) on how your writing went.

Benefits of Daily Writing

Writing daily helps align your time with your evaluation criteria (e.g. 80% research and 20% clinical).

Table 0.4: Benefits of daily writing. Adapted from NCFDD {tbl-benefits-daily-writing}

Benefit	Description	
Productivity Leads to slow, steady productivity and fewer		
Shift	feelings of anxiety over meeting writing	
	expectations	
Mental	Writing is the most important part of your	
Shift	success; therefore, it is your top priority.	
Behavior	You write everyday and find a way to be	
Shift	accountable that works for you!	

Academic Medicine Jobs

AAMC Faculty Salary Report

Looking to get an idea of academic faculty salaries? The annual AAMC Faculty Salary Report compiles academic faculty salaries by rank, degree, department/specialty, medical school

type, region, and more. This is often available for free through your university library. Get to know your librarian!

Tenure-Track Offer Letters

What goes into a tenure-track offer letter? The Burroughs Well-come Fund provides a comprehensive list of offer letter components in their article, "Academic Tenure-Track Offer Letters."

NIH Loan Repayment Program

NIH Loan Repayment Program

Online Resources

Edge for Scholars

Edge for Scholars

NCFDD

NCFDD

Professional Organizations

American Physician Scientists Association

Suggested Readings

Not Discussed

Not Discussed by Michael Stein

Publishing Your Medical Research

Publishing Your Medical Research by Daniel W. Byrne

References

- Jia, Xiaoming, Buhm Han, Suna Onengut-Gumuscu, Wei-Min Chen, Patrick J. Concannon, Stephen S. Rich, Soumya Raychaudhuri, and Paul I. W. De Bakker. 2013. "Imputing Amino Acid Polymorphisms in Human Leukocyte Antigens." Edited by Jianming Tang. PLoS ONE 8 (6): e64683. https://doi.org/10.1371/journal.pone.0064683.
- Viard, Mathias, Colm O'hUigin, Yuko Yuki, Arman A. Bashirova, David R. Collins, Jonathan M. Urbach, Steven Wolinsky, et al. 2024. "Impact of HLA Class I Functional Divergence on HIV Control." *Science* 383 (6680): 319–25. https://doi.org/10.1126/science.adk0777.
- Zheng, X, J Shen, C Cox, J C Wakefield, M G Ehm, M R Nelson, and B S Weir. 2014. "HIBAG—HLA Genotype Imputation with Attribute Bagging." *The Pharmacogenomics Journal* 14 (2): 192–200. https://doi.org/10.1038/tpj.2013.18.