

Monday - 10 January 2022

Today was my second time setting foot into Biopolis, GIS to be specific. The first being yesterday when I visited to familiarise myself with my new working environment. This internship is my first internship and I was not too sure what to expect. However, listening in on a meeting with Jing Mei, Peh Joo and Zi Lin in the morning helped me better integrate into the environment as I slowly learnt about the method of operations and also to set some expectations. Their projects and discussion brought some new perspectives and helped me better understand the innovation of their projects. However, there were some troubles with the Internet throughout the day, as I struggled to properly log into the desktop and laptop at level 6. I had to spend some time at level 8 to complete a few of my CITI courses but thankfully, IT help came through at 3.30 and I was finally able to connect at level 6. Laptop started lagging some time later and had to initiate updates and to restart the laptop. Hopefully, I'll be able to complete the CITI courses as soon as possible.

The part I enjoyed the most today was the brainstorming. I liked that the ideas that were bouncing off each other during brainstorming was very fresh and thought-provoking.

Tuesday - 11 January 2022

Today's schedule was to work from home in the morning and return to office in the afternoon to sort out my VPN so that I can start working from home too. Listened in on Zi Lin's meeting this morning and was exposed to many forms of statistical analysis, some of which I have never heard of before (e.g. Kaplan Meier). Attempted to come up with an explanation for the results on my own but it was relatively difficult to arrive at a reasonable/logical sound conclusion. Felt a little lost initially because I am still trying to grasp and understand the different projects that are going on. Finally completed my CITI today.

Zipped the files that Peh Joo had already generated and sent them out to the correspondents for the project. Managed to log into R drive using one of the student desktops available at level 6 and Peh Joo taught me how to run the program using R. Was very difficult to understand what was going on because I have very little background in programming (learnt very basic python previously) and had a hard time understanding. Peh Joo was very understanding and helpful in trying to help me make sense of the codes and to also let me know what the important codes are. Attempted to re-run the code on my own but faced a little trouble as the desktop did not have a lot of libraries installed. Finally managed to run it successfully after some help from Peh Joo.

Will be reading up on Zi Lin, Peh Joo and my project proposals to get a better understanding of what each project entails.

Wednesday - 12 January 2022

This morning, I joined in on Jon's project meeting. Was a little lost and confused with regards to the focus group coding. I think Jon's project was slightly more familiar to me with regards to the aspect that is being investigated. The idea of investigating mammogram uptake in older women through the lenses of their adult children was very fresh and interesting. Alongside the results from Zi Lin's project, I was inspired to want to know more about whether a person's level of education/generation affects their willingness to go for annual/biennial screenings.

From the results obtained from their project, it seems that some of the reasons hindering older women from actively going for mammograms is the lack of appropriate cues and fear of the unknown. As opposed to the adult children who are more exposed to technology, the older women have more fear and uncertainty towards mammograms, and they are unable to alleviate their fears. They are only able to gain more information through hearsays and sometimes this information is passed on inaccurately. Additionally, there may not be sufficient emphasis placed on the importance of mammography and breast cancer. The lack of awareness of the severity of breast cancer may result in them not feeling the need to go for mammography.

I finished reading Jon, Zi Lin, Peh Joo manuscripts and the proposal on gut microbiome changes in chemotherapy-treated patients. I had to do additional readings on the different types of statistical analyses that were used because some of them were unheard of to me. For example, I learnt that Cohen's kappa is usually used to assess the agreement between 2 raters.

Besides that, going through the proposal in detail prompted me to clearly understand the goals and objectives of the project and to also think of the directions that this project could go in. Reading through the proposal, I also learnt that there are many aspects to a proposal and besides the experiment and results alone, there are also other details to pay attention to in the proposal. For example, mentioning the team's strong track record of recruiting patients and healthy individuals for various studies in SG to increase the credibility of the study.

Thursday - 13 January 2022

Today I joined in on the SST HTPO Townhall in the morning to listen to the different projects conducted by the Seed Fund Recipients of 2021. I was very impressed by the slides prepared by the presenters. The slides were very informative and very concise. In the afternoon, I read through Peh Joo's updated research proposal and tried to think of pointers from her perspective. I think engaging different people to proofread/read through new works can help tremendously in terms of providing constructive criticism as to whether the points mentioned in the paragraph is what the author is trying to convey, or to offer new points of view. I realised I am lacking in the knowledge department when I attempted to think of new pointers and to also counter them successfully. I realised that when data is given and I try to think of a plausible explanation, I always fail to logically reason it out. Reading more related/interdisciplinary works may be helpful to me for a head start.

Friday - 14 January 2022

This morning, I joined in on Zi Lin's sharing on her project. I realised that as much as I understand the manuscript, whenever it comes to the analysis of results/data, I am always very confused. I feel that if I were to be the one receiving the results, I would have absolutely no idea where to even begin, i.e. sorting the data, categorising them properly, and much less to even begin to make sense of it. For example, I may make the mistake of categorising them by tumour size instead of age etc. Wrong categorisation of data may result in misinterpretation and as a result not yield any results for the factor investigated. I think I need more understanding with regards to tests to yield data.

I also joined in on Jon's project meeting with Sheen to learn more about coding for his project and to use Nvivo. Zi Lin was very nice to guide me along and let me know what was going on. I learnt that coding can be applied to qualitative, not just quantitative analysis of results. I think I am still struggling to understand the coding part but I think I am beginning to get the hang of what to expect and what I am working towards for the coding part. I am less lost (though still lost) for it.

I also joined in on the weekly Friday seminars to learn more about transcriptional genome though I do not really understand it thoroughly.

Today concludes my first week of this internship. I am very humbled by the projects and the experience and I hope to learn more even though I sometimes feel that I have nothing much to contribute.

Monday - 17 January 2022

Today I focused on trying to understand the code package for BREATHE. I took a course on python in year 1 but I struggled with it. Having to run the scripts using R is a bit difficult for me, especially because I do not have a very strong coding background. I mostly remember the part about having to ensure that the names of the package/file in the code matches that in the R drive folders. Other than that, I am trying to study the codes that are not within the input section. I am still trying to understand the codes and how Peh Joo managed to come up with these codes from scratch, especially because she has different codes for different categories.

Tuesday - 18 January 2022

This morning, I joined in on Zi Lin's project meeting. Everytime I join the meeting discussion, I realise that the results/data may not always match the hypothesis and that there may not always be an obvious plausible reason to explain the data. Sometimes, the explanations are not as straightforward as expected and there may be other reasons behind the discrepancy.

I also continued studying the code for BREATHE because other than the coding aspect, there is also the statistical aspect in the code. If I am unable to choose the correct statistical test/analysis to conduct, I cannot come up with a code that will be useful. I think I need to not only brush up on my coding knowledge, I need to also look up the commonly used statistical tests and analysis used in these projects.

I hope to be able to understand the code and attempt to run it successfully tomorrow.

Wednesday - 19 January 2022

I read several articles regarding breast cancer screening in western countries and the types of breast examinations conducted (BSE, CBE and Mammogram), and the take up rate of cancer screening in those countries. Those countries have similar profiles to high-income Asia countries but I think the success of breast cancer screening in those countries (compared to high-income Asia countries) might be due to the fact that they are more liberal and it is less of a taboo to speak about it.

Today I focused on reading more on the policies/screening methods in different countries and the possible factors contributing to its success. Tomorrow, I'll be reading more on breast cancer screening in Asia countries and the differences in take up rate in high-income countries and low-income countries.

Thursday - 20 January 2022

Today I found several articles regarding breast cancer in high-income and low-income Asia countries. I realised that there are different reasons deterring women from screening for cancer. In higher income countries, the factors deterring women from participating in breast cancer screening could be due to the lack of emphasis or fear of diagnosis. However, in lower income countries, it is very common that most people have not even heard of mammogram and much less attend annual or biennial screenings. There is also insufficient efforts to setup a nationwide screening to encourage/promote cancer screening and spread/increase awareness. I have consolidated some data on the uptake rate and organizations set up in the countries that are responsible for promoting cancer screening.

Tomorrow, I'll start putting the data into a table for the Asian countries.

Friday - 21 January 2022

After looking through the papers for the different countries, I started working on a table today. I thought the data collation was the tougher part as the study conducted in the different papers had their own set of criterion. This means that the exact data that I wish to add into the table may sometimes be unavailable. I think being able to refine my searches will be a good skill to acquire.

Monday - 24 January 2022

I focused on creating the table for the list of Asian countries today. I think the thing about compiling data from various sources is consistency. Sometimes the data obtained from a paper may be similar to what I am searching for, but may not always be the case. E.g. while searching for the screening age, some published works may quote 40 years, while some may quote 45 years. I think the source from which I acquire my data has to be verified by checking with the national organisation's website.

Some data from my searches may be outdated. Again, I need to work on refining my searches on search engines to get the most recent & accurate data.

Tuesday - 25 January 2022

Today's focus was mainly lit review and finding data from various sources to put into the table. I think literature reviews are important but it can take a while.

Wednesday - 26 January 2022

Today's focus was also lit reviews, but during check in, a few other factors can be considered too. I.e. gini coefficient, human development index. During check in, I think I had a realization that the factors deterring women from receiving or participating in screening could be vastly different even if the take up rate is equally low in different countries. Low take up rate in western asia/middle east could be due to stigma/patriarchal system deters them from seeking help while that in LMIC SEA countries could be attributed to lack of access to healthcare etc. I think this is a reminder that I have to consider factors from various angles when working on a project.

Thursday - 27 January 2022

This morning, I joined in on Peh Joo's work ability questionnaire meeting. I think it was very interesting to hear the different takes/perspectives from everyone during the meeting. I enjoyed learning about crafting specific questions to obtain necessary information. I realised even tweaking the question slightly will result in different types of information obtained. During the meeting, I learnt that it was very important to ask the purpose of each question and if there was a point to those questions.

While doing lit reviews, I found published articles that are similar to what I am working on. I think it is a good reminder that literature review is always about finding out new things, and not just finding information I already know.

Friday - 28 January 2022

Today, I found a published book that has information that is very comprehensive and similar to the data I have been searching for. However, the book provides only certain data and does not include other factors such as GINI for comparison.

I also learnt how to merge files on excel today using R!! I learnt the merge function and changing of column names. I think R is not as difficult and daunting as I imagined it to be, though it is still difficult. After coding successfully once, I think I am slightly more confident to try. I have also found out more functions from Peh Joo so hopefully I can incorporate them into the actual code.

```
0 2 4 6 8 10 12 14 16 18 20 22 24 26
#CHANGE DIRECTORY TO WHERE FILES ARE
setwd("Downloads")

#READ MAMMOGRAPHY INFO IN
data <- read.csv("DEVICES08.DEVICES09.DEVICES10.DEVICES11.DEVICES12.DEVICES21.DEVICES22.DEVICES23.csv",header=T)

#GET NUMBER OF ROWS AND COLUMNS FOR FILE READ IN
dim(data)

#ORDER DATA BY COUNTRY FIRST, THEN YEAR, IN DECREASING ORDER
data <- data[order(data$Country,data$Year,decreasing=T),]

#CREATE NEW OBJECT FROM DATA 1, ONLY ONE ROW KEPT FOR EACH COUNTRY (LATEST DATA)
data1 <- data[!duplicated(data$Country),]

#READ IN NEW FILE, (header is true by default. indicate header=F if first row is a data)
gini <- read.csv("GINI.csv",header=T)
gini <- gini[,c(1,51)]
colnames(gini)[1] <- "Country"

# CREATE ASIAN DATASET EG ASR, (all.x=T) means the countries will still be printed even if there is no data. will appear as N/A
#E.G. [,] IS LIKE X & Y, AND WHICH COLUMN/ROW TO SELECT ETC.
merged <- merge(asia,data1,by="Country",all.x=T)
merged <- merge(merged,gini,by="Country",all.x=T)
merged <- merged[,c("Country","Year","Total density per million females aged from 50 to 69 years old..Mammography.units","GINI.2014")]

#CHOOSING THE COLUMNS. eg specifically those 2 columns
yh<-yh[c(1,60)]
```

Monday - 31 January 2022

Today, I tried merging the files and exporting them, but somehow it couldn't be properly exported into excel, and the data that is supposed to split into 4 columns was merged into 1. Tried many methods and got it to work, though I am not too sure what went wrong.

Thursday - 3 Feb 2022

Found a few indexes that I thought could be useful. Merged the indexes but there were a few mistakes. I think the mistakes appeared during merging because the countries are named differently, depending on the source. Some countries like South Korea and North Korea are named differently in IARC's data bank. I also learnt more on dataset management today! Previously, I only referred to the coding functions and did the merging of datasets without keeping the codes and creating many many intermediate files. I had no way of reproducing the exact same dataset because I did not keep a record of the codes and I also made changes to the excel itself, instead of doing it on R.

Friday - 4 February 2022

Managed to come up with table 1 that has quite a bit of info. Took quite some time to understand the code that was used for cleaning the GINI dataset. I think that code is very useful in cleaning the datasets and can be used for other datasets in the future. Also spent some time looking at the newly created table 1 for information. Will be looking into linear regression models and PCA and other potential forms of analysis.

```
for (i in 1:nrow(gini)){ #start of a loop. i increases from 1 to no. of rows in gini
temp <- gini[gini$Country.Name==gini$Country.Name[i],] #temp is a name. naming. so gini$country.name is the
new name for gini$country.name[i]
temp1 <- as.data.frame(t(temp[,5:65])) #t is for transposition, changing columns to rows. columns 5 to 65.
as.data.frame is to convert temp to data frame. it was previously an array
temp1$Year <- rownames(temp1) #converting rownames to header called year
temp1$Year <- sub("X","",temp1$Year) #replace the X in "temp1$Year with space. basically to remove the X
temp1 <- temp1[is.na(temp1[,1])==F,] #removing data in temp1 column 1 that contains NA. **RIGHT NOW, the
data is only being run for the first country in the dataset. so all the NA are removed and only cells with data are
retained
temp1 <- temp1[order(temp1[,2],decreasing=T),] #change the order of the data such that the most recent data is
at the top
gini.m[i,1] <- gini$Country.Code[i] #put country code into the table
gini.m[i,2] <- paste0(temp1[1,1],"(",temp1[1,2],")") #paste data for each individual i into the table
}
```

Monday - 7 February 2022

Struggled in the morning to code for statistical models to use for my data. Finally managed to squeeze out a linear model but I have no idea what the numbers on the screen mean. Stared at the numbers for a little while and after some help from Zi Lin & Peh Joo, I can somewhat understand the numbers. They've also offered some tips on how to do boxplot, scatterplot and barplot. Will be fitting different numbers tomorrow to see if I can find any useful information from the data & maybe some relationship! I think I'll be spending some time on the generation and interpretation of data.

Tuesday - 8 February 2022

Managed to come up with a few association plots, using ASIR and ASMR. However, I was unable to get any plots using Number of Physicians and Number of mammogram units (continuous data). I later found out that it is because the data is value (year) and R was unable to read the information. Spent some time learning how to split the characters, but was stuck halfway. Zi Lin came through and taught me how to process the split data.

Wednesday - 9 February 2022

Today's focus was to analyse the figures obtained and suggest plausible explanations to the figures and trends. I think what I've read during literature reviews came in handy. However, sometimes I think I get confused by the factors, i.e. when drawing conclusions from ASMR and ASIR, I might get a bit confused with whether my explanation is prevalence-related. I think it's important to me to always take a step back and see the bigger picture after each analysis. The figures I've obtained from R are still quite simple, perhaps not drawing out the potential r/s and trends. Will try to see what other figures I can obtain tomorrow.

Thursday - 10 February 2022

This morning, I learnt how to come up with a figure with more variables using ggplot (with a lot of help from Peh Joo). As compared to the previous figures that I came up with on my own, this plot has more variables and more information. It is easier to understand and digest the data at one glance. Other than ASIR and ASMR, I managed to include income groups and regions using colours and symbols into the figure. After fumbling around for a while, I managed to also include the number of mammography units as size into the figure! yAY!!

Will be working on expanding the table to include more relevant information such as breast cancer take up rate in countries with national programs and see what can develop from there.

Friday - 11 February 2022

Haha turns out the figure I generated yesterday (using size as a variable) was not entirely correct. Managed to correct it today. The function `scale_size_continuous` was used to scale the variable's range (0 to 400) to (0, 10). Also tried a new task today to use R to randomly generate numbers for toto and even included the sound in.

Sat in on the meeting updates for Seeu Si and Jonathon's project. Very interesting to be involved in the brainstorming process of the recruitment of participants and also to learn more about systematic reviews from Seeu Si's sharing on IGM. Learnt a lot about the challenges of using a forum to collect responses from participants and recruiting from polyclinics etc.

Refined the figure by including NA countries into the figure by creating a new variable in the dataset. Will be working on adding an asterisk to selected countries with no information for number of mammography units in the figure.

Monday - 14 February 2022

Attempted to add an asterisk to the countries with N/A for the figures but faced a bit of difficulty. Tried using loop function, mutate and if/else but it did not really work :(Will try again tomorrow :/

In the meantime, I started working on table 2, which is essentially an upgraded version of the very first table I was working on. While working on table 2, I realised that the initial version of the table I created was lacking in terms of details. Some details were grouped together and ended up taking on a different meaning. Included more details and removed columns that have already appeared in table 1.

While coming up with the new table 2, I also realised that some of the data in the published book is outdated. Will have to double check the data to make sure that the numbers given are the most recent/updated.

Tuesday - 15 February 2022

Continued working on table 2 and had to reconcile the differences between the numbers I initially obtained for the first table and the data in the published book. Included the doi in the table and it looks more organized. I think compared to the first table I came up with, table 2 looks more informative and comprehensive despite having less columns. Will be attempting to finish table 2 tomorrow and try out the side quest for this week!

Wednesday - 16 February 2022

Today, I tried adding an asterisk to the country code in the figure. However, I realised the results vary a lot. Sometimes, the figure will appear nicely but other times the symbols will disappear. I don't know if it is because sometimes when I edit the code to try to get it to run, I accidentally alter some parts of the code instead. I spent some time relooking the code and correcting it. The code now runs and produces the same figure each time I run! Note to self: don't edit the code directly

Thursday - 17 February 2022

This morning I managed to add the asterisk to the country code in the figure plot and also adjusted the sizes of the plots and country code. I think as a whole, the figure is very informative. Despite having quite a few variables in the figure, it looks very clean and neat. I am also almost done with table 2. After this afternoon's check in, I will be working on coming up with a figure legend and splitting table 2 into 4 separate tables according to the type of screening program available in those countries. I think splitting table 2 into 4 tables might be better as it will look clearer and the countries can be better distinguished by the program type to help supplement the figure.

Friday - 18 February 2022

When reading up on breast screening in Vietnam, I came across this particular article that talked about systematic review. Prior to last Friday's meeting, I did not know what systematic review was. After Seeu Si's sharing on her IGM project and how she conducted her systematic review and some explanation from ZiLin, I think it helped me better understand the article. Other than just obtaining data, I think conducting literature reviews has helped me learn a lot more about scientific research in general.

Monday - 21 February 2022

Worked on the table today. Almost finishing the table with a few countries that have limited information online. The information available online for certain countries with no population-based screening/established opportunistic screening is usually not accurate as the information varies from sources to sources.

Tuesday - 22 February 2022

Used R to split the table 2 into 4 different tables – countries with population-based mammography screening, countries with population-based screening (not mammography, i.e. BSE/CBE), countries with established mammographic screening (even though it is opportunistic, it is still widely accessible and available to the general public), and lastly, countries with little to opportunistic screening (screening only offered occasionally, not stationed at a particular location/only offered in the capital city/only with one available machine for the whole capital city etc)

I think previously, I would have done everything (reclassify the tables into the 4 other tables) manually but I think now I am moving towards trying to use R more. I learnt a new function (`%>% pipe` today). Not entirely sure what it does, but I vaguely know when to use. I think when the time comes for me to have to use this function I can try again.

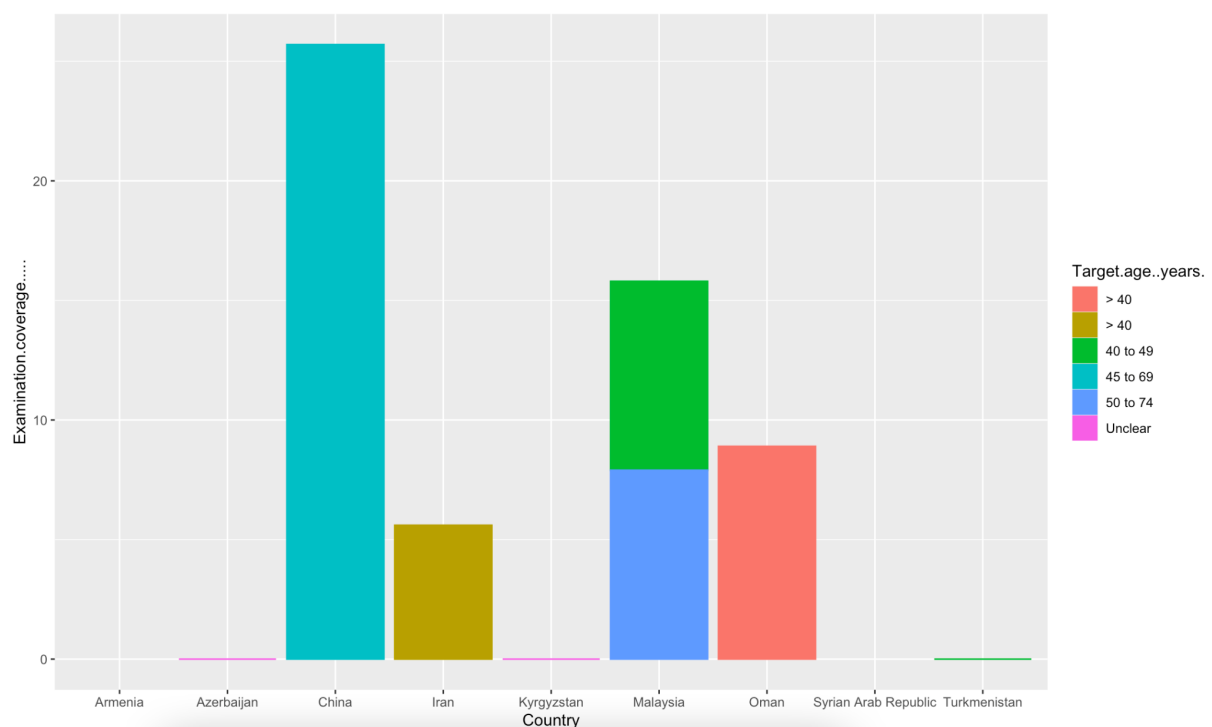
Wednesday - 23 February 2022

Wanted to help send the risk reports for BREATHE but was unable to password protect the files using the work laptop despite asking Peh Joo for help. There isn't the proper software on the laptop for that. Shall go to office tomorrow to use the work desktop to do it. Tried plotting a barplot for the 4 different tables, with countries and mammographic take up rate but there was an error. Since different countries have different screening methods, some countries may appear twice, which resulted in me not being able to create the barplot due to duplicates in the x variables.

Thursday - 24 February 2022

YAY! Managed to come up with the barplot on my own. Was close to giving up and asking for help but I tried several methods and one of them worked. Tried using the `make.unique` function but it didn't work. I tried using `ggplot` afterwards and it actually worked. I also managed to add in colours to the barplot to include more information. I was playing around with the variable to add in but I am not sure if I should include the type of screening practice or the interval, because 2 of the tables are for MAM only, so using screening practice as a variable is not very useful. Initially, the colour of the barplot was only the outline but after some google searches, I managed to fill in the colours too.

One of the plots I managed to come up with, using target age as a variable



Monday - 28 February 2022

Did more literature review today! Focused on breast screening in European countries. A very large proportion of European countries have very high breast screening (MAM) as compared to Asian countries. I think the main reasons as to why the MAM screening rate is much higher in European countries could be because Western countries are more liberal in terms of their culture. As a result, there is less stigma and taboo surrounding breasts and screening.

Tuesday - 1 March 2022

I found a very interesting article today!! The article reviews all the screening programs in European countries and another study compared neighbouring countries with similar demographic profiles in European countries. This might be applicable to Asian countries with similar profiles, such as Cambodia and Laos, or Hong Kong and Taiwan.

Wednesday - 2 March 2022

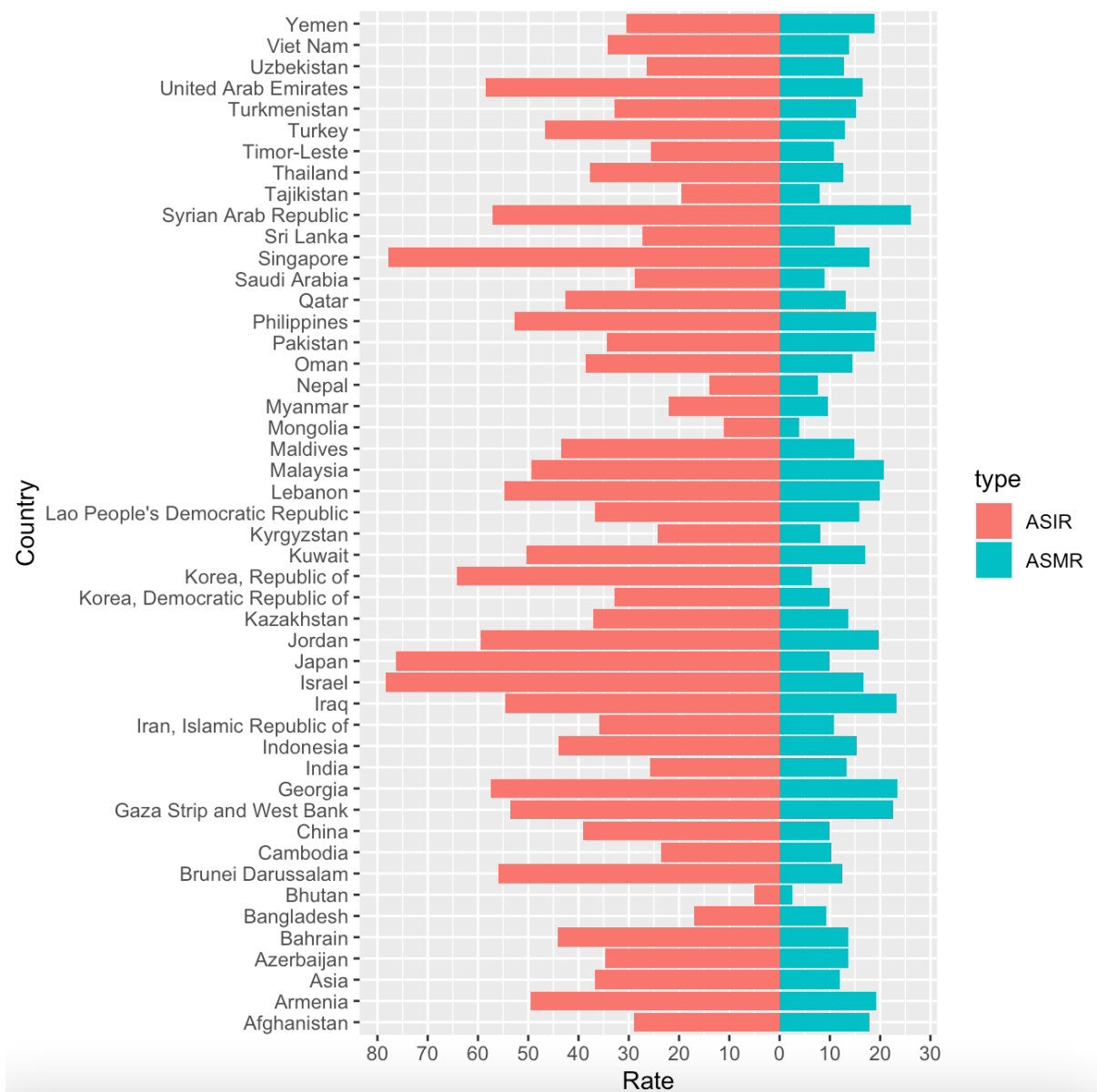
Adjusted the code for the figures. Previously the figures were not accurately representing the data obtained as it was stacked (total). Adjusted the code to separate the bars such that the screening rate was not the total sum, but rather, just the percentage of screening for each practice. Also, continued doing literature reviews for screening in European countries. I think the Nordic countries have been fairly successful in their screening programs, and if Asian countries want to increase their MAM screening take up rate, they should consider emulating that of Nordic countries and alter it to fit better into their culture/practices.

Thursday - 3 March 2022

Tried writing the abstract and introduction for the review. I think there is still room to add in a lot more information and to also make it more scientific, and less narrative. The points that are currently written are very similar and the ideas overlap a lot. I think I will need to focus on distinguishing them and to also add in different ideas.

Friday - 4 March 2022

Presented the review briefly during the weekly roundtable meeting. Was a little nervous but everyone had a lot of nice feedback and contributions!! Felt daunted to be given a task to create a population pyramid using R that is as informative & nice as the sample article, but was given a lot of tips and ideas on how to create it by everyone at the table. Tried to create the figure subsequently and struggled the whole afternoon to create something that is remotely similar to that of the sample article HAHAAHA. I think it was very challenging, but fun nonetheless. Asked Zi Lin for help and finally managed to come up with a decent looking figure after a couple of hours HAHAAHA



Will have to work on splitting the countries by region, income and indicate the #MAM at the side.

Monday - 7 March 2022

Analysed the code for the new figure 1 today. I think compared to the previous figure 1, this figure is more informative. Although both figures had the same information, the trends and numbers are easier to see and understand at first glance in this new diagram. I think the difference is that the region is now classified by space and colour intensity rather than shapes and colours. There are also more numbers in this new figure which gives the reader a better idea. Also, been reading on Europe and Asia and looking for similarities.

Tuesday - 8 March 2022

Morning updates. Learnt more about NCAM and the conference that the review is (most likely) going to be presented at. Quite nervous about having to ultimately finish the review and maybe present it in front of a crowd, but at the same time, I think that it is an amazing opportunity to gain an experience like this. I think the standard and requirement of presentations and reports I have written in school is incomparable to the standard for this review but I do think that I can make use of what I have learnt and use this opportunity to hone my skills.

Wednesday - 9 March 2022

Not sure if I was dreaming AHAHHA but somehow I couldn't find the ASIR for both Asia and Europe in GLOBOCAN. Maybe the information really did not exist and I happened to read it elsewhere. Spent my morning trying to find that information on GLOBOCAN. Also, managed to find some more points to include in the introduction for risk factors. I think the numbers for risk factors are very interesting and can offer different views. BRCA1/2 mutations, modifiable risk factors etc all contribute in varying extents to breast cancer. I have always thought BRCA1/2 mutations will contribute significantly to breast cancer, but interestingly, it only contributes about 5 to 10%.

Thursday - 10 March 2022

Tried putting the points into the introduction but many points overlap. Introduction seems kinda repetitive and lengthy now and many points are poorly paraphrased. Relooked into mammography to consider its effectiveness/usefulness. I think previously my points for mammography were slightly biased, focusing only on the good, without considering the bad. The previous recommendation of bringing the screening age forward is only applicable if there were no associated harms. Have to reconsider screening in Asia as a whole.

Friday - 11 March 2022

Morning meeting helped to clarify some of my doubts because there were a few things to consider for the points, i.e. describing all types of mammography etc.. Came up with a list of things to do. Will be coming up with headers and points for each section in the introduction. Also tried to think about why Asians have an earlier peak age for incidence but can't seem to find much information on that. Had to bounce some ideas off Zi Lin and look for more information online. Will look into other factors such as age at menarche, BMI at 14 years old etc.

Monday - 14 March 2022

Managed to come up with the first paragraph for the intro! I think some work is still needed, but I think it's clearer than before and there is some sort of order to it, rather than just throwing the points around. I think the ASIR description from the genre article is nice! Tried to find similar data for more recent years.

Tuesday - 15 March 2022

Managed to come up with the second part of the intro today! I think it was more coherent and the points do not overlap as much anymore. Previously, I was struggling a lot to put the points together but I think some of the points were not very useful, so I tried weeding them out. Managed to focus on screening in Asia today.

Wednesday - 16 March 2022

Met Alya on zoom today! Introduced some things that I struggled with when I first joined and talked a bit about projects and expectations. Also managed to finish a small part of the Introduction, on the cons of mammography screening. While writing on overdiagnosis from mammography screening in Asia, I realised that there is very little information in Asia. While overdiagnosis is one of the main factors deterring people from obtaining mammography screening, little is studied about it in Asia, maybe because mammography screening is much less common in Asia as compared to European countries.

Thursday - 17 March 2022

Finally finished the other con of mammography screening, which is the effect of false positives. Similar to overdiagnosis, I feel that there is too little information on detailed mammography screening. It seems like there is a lot yet not enough information.

Found a few articles that contradict each other i.e. some articles saying there is little effect of false positives on take up rate, some other articles saying that false positives is enough to set people back from mammography screening by up to 24 months. I think at this point, it is also important to note the particular countries that presented these data because it may not be the same across all European countries as it is across all Asian countries. Also, had a short meeting with Peh Joo and Alya for the chemobiome project to set the expectations and timeline. Very excited to embark on the chemobiome project!!

Friday - 18 March 2022

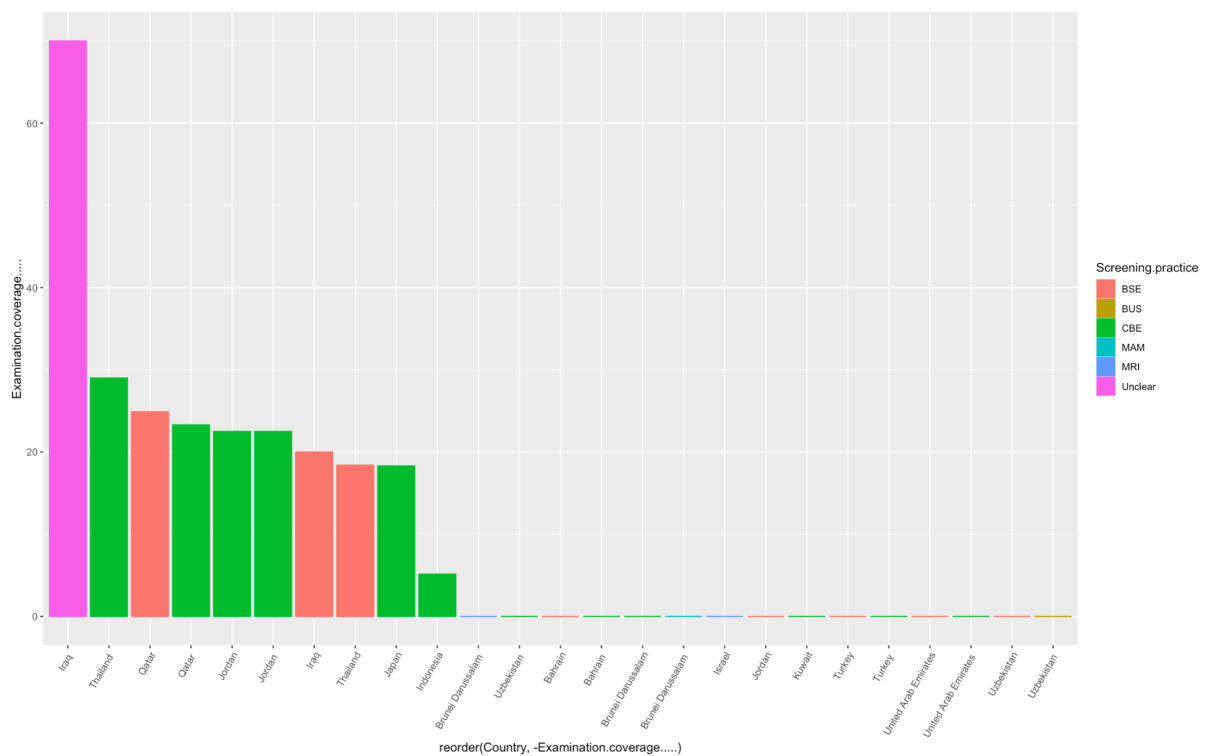
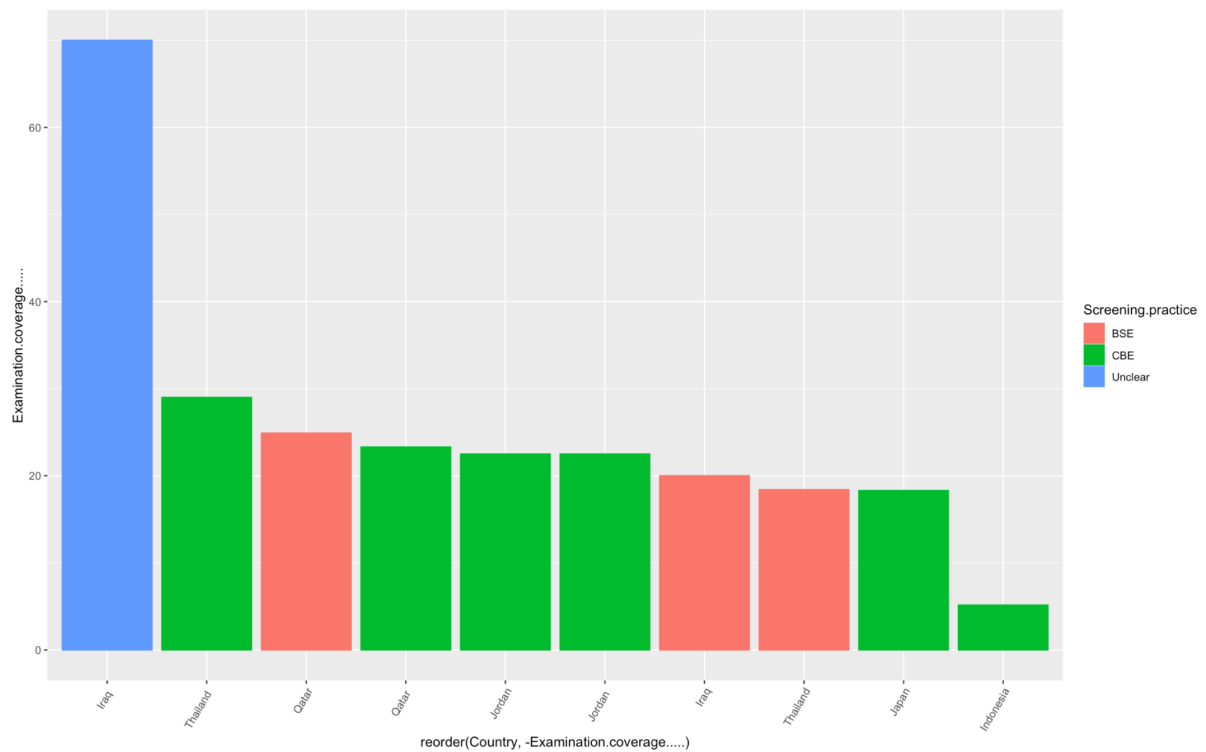
Met with boss and managed to iron some things out. I think I was very confused by all the contradicting articles and also of the lack of content because of the lack of information available in Asia. Will focus on cons of mammography in Asia, specifically the overdiagnosis and false positives in countries in Asia. Will just look for information in all countries and not just a general trend in Asia, since there isn't enough/sufficient information. Will also redo the diagram, just slightly different from the ones created previously.

Monday - 21 March 2022

Recoded the figures for population based mammography screening. Had some problems with the descending order but managed to figure it out. The figure is still lacking in terms of the order and colour.

Tuesday - 22 March 2022

Coded for the other figure, managed to come up with 2 figures. One with the countries that has unclear information and another cleaner version with lesser clutter.



Wednesday - 23 March 2022

Attempted the side quest and tried to calculate the change in ranks. Managed to consolidate all the necessary data into one data set, but struggled to calculate the change for ranks that are in a certain range. I tried to turn the ranges into absolute values but it is not working out too well.

Thursday - 24 March 2022

Looked for cost effectiveness analysis and number of deaths averted per country as part of the evaluation of effectiveness of mammography screening in Asia.

Surprisingly, there are quite a number of articles that point towards mammography being not cost effective in some countries. I think one thing I learnt from this review is to go in the direction of the evidence rather than looking for specific articles, though not cherry picking, but it may not be the most accurate/holistic review.

Friday - 25 March 2022

Focused on looking for evidence and analyses on the various countries today. I think I was lucky to have chanced upon the article on why mammography screening was abandoned in China when I was looking for other articles previously. Due to the lack of mammography screening programs available, there is fairly little evidence/article on how false positives/overdiagnosis may hinder future screenings. I think rather than trying to find evidence of something that may or may not exist, I can try to channel my search towards available resources, such as the QALYs and ICER of mammography screening in Asia.

Monday - 28 March 2022

Managed to refine the false positive and overdiagnosis in Asia portion today using the information I obtained last week. I think adding in numbers obtained from Asian countries helps to highlight the concern of mammography screening - or the lack thereof - in certain countries. The numbers obtained for the number of false positive cases for each true cancer identified in Asian countries and Western countries vary greatly when I was doing the search and I realised this too is a problem. Will work on cost effectiveness analysis and the breakdown of it in each country tomorrow.

Tuesday - 29 March 2022

I spent the better half of this morning reading on ICER and cost effectiveness analysis. I think prior to using CEA, I have never really invested much time in reading about health economics. I was introduced to a few new terms such as cost benefit analysis etc. I focused on the components of ICER and CEA today to discuss the various countries' stand on whether it is cost effective to have mammography screening in the countries.

Wednesday - 30 March 2022

Came up with the cost effectiveness analysis draft today. Also tried coming up with future directions for the review. I think moving forward, rather than attempting to change the entire mammography screening because of its limitations, the future direction it should take should look into filling the gaps and improving the current system until a better screening tool is developed.

Thursday - 31 March 2022

Will be working on ways to mitigate the problems of mammography screening and how to improve it. I think to effectively fill the gaps, the problems have to be identified, previous solutions used to solve the problem, the flaw with those solutions, etc. Will also start looking into translation of questionnaires for the chemobiome project! Am fairly excited to start on that

Friday - 1 April 2022

Started looking into ways to mitigate the problems of mammography, i.e. complementing mammography with tomosynthesis, introducing CBE/BSE prior to mammography to increase cost effectiveness etc..

Monday - 4 April 2022

Introduced sub headers and topic sentences into some of my sections today. I think right now, the sections and subsections are still a little disorganised. The content for each part overlaps a bit and I'm stuck in some sort of catch-22 HAHA.

Tuesday - 5 April 2022

Had a meeting on workplace harassment and did a bit more research on the types of harassment and resources available for each type of harassment. Did literature reviews to look at how to conquer the problem of overdiagnosis and false positives.

Wednesday - 6 April 2022

Still not very successful in finding significant data for overdiagnosis and false positive rates in Asia or just in general. Reorganised and worked on the overdiagnosis part. I think for overdiagnosis, there are a few points that are similar yet different, so I am working on trying to be more explicit and concise in that section.

Thursday - 7 April 2022

Felt unwell and worked from home. Did literature reviews on screening programs and guidelines in Asia. Managed to find some new information to support the content but some parts of the overdiagnosis and false positive parts are still lacking.

Friday - 8 April 2022

Attended a webinar on manuscript writing in the morning. Had a mini breakthrough. Asked Zi Lin and Alya for their opinions and managed to restructure and come up with a skeleton for the review. I think going with the new skeleton, the review will have a nicer flow and organisation to it. Will use the backbone for restructuring next Monday!

Monday - 11 April 2022

Attended a meeting in the afternoon on coding & transcribing for Zi Lin & Jon's proj. I think it was quite fun to watch the recorded FGD and also to learn more about Su-Ann's take on the project and how she can contribute. Also managed to meet boss to clean and pre-finalize the first part of the review. Also learnt about the importance of oxford commas!! HAHAHAH. I learnt English and grammar in general without oxford commas and have always thought that they were weird HAHAHA but I think the example made sense!!

E.g. she was wearing a big, black and blue hat
Vs she was wearing a big, black, and blue hat

In the one w/o oxford comma, the sentence can be interpreted as the girl wearing a big bruised (black & blue) hat.

Tuesday - 12 April 2022

Worked on the changes we discussed during the meeting yesterday. I think there's a lot to improve on in the writing area. I think the information I have obtained are relevant, but somewhat repetitive. In terms of writing, I think the points I have made do not provide sufficient insight and is not concise enough.

Wednesday - 13 April 2022

Had another check in this afternoon and reshuffled some points to better describe the screening scene in asia. Several things to work on:

- 1) Check that asia and europe refer to the same countries throughout the paper
- 2) Replace the source's link -- always use the direct source, not the factsheet/quote other journals
- 3) Be more explicit in writing, i.e. higher asmr = more breast cancer deaths
- 4) Re-write M/I ratio. E.g. instead of 0.26, change to 1 in 5 breast cancer cases die
- 5) Consider including the U.S. when comparing to western countries
- 6) Use updated numbers if possible
- 7) Remove the focus on breast cancer problems/risk factors etc. and focus on mammography screening in asia and its problems/barriers

Thursday - 14 April 2022

Presented the introduction to the team during roundtable meeting this morning. Went for GIS carnival yay!!! Was very fun to see and experience a semblance of pre-covid life.

Monday - 18 April 2022

Had a check in this afternoon. Discussed about the flow and structure for intro. I realise having a good flow and getting straight to the point increases the impact and readability of the article. Having good headers and structures also sets the direction of the article as early as possible. As usual, stuck in a catch-22 -- not sure whether inequities in breast cancer outcomes or income should come first. I feel like there is something to unlock the relationship between both and generally throughout the review but I can't seem to figure out what is lacking. Not sure if it is because the points made in the first place are so jumbled up which is why I'm struggling to see a clear flow/structure.

Tuesday - 19 April 2022

Worked on linking inequity in income and breast cancer outcomes. I think a clear link has to be established in order to effectively communicate the point and make the jump less abrupt.

Wednesday - 20 April 2022

Worked on mammography screening, the organisation and additional info such as stage at presentation. This is inspired by Yip Cheng Har's paper, where she presented ASIR and ASMR by age, stage and trend. Had to look for more information on stage at diagnosis, 5 year survival rate by income groups

Signposting words

Contrast

- However
- Whereas
- On the other hand
- Nevertheless
- Although
- Yet
- Despite
- In contrast to
- By contrast

Similar

- Likewise
- Similarly
- Also
- As well

Result

- Therefore
- Consequently
- Thus
- As a result (of)
- Due to
- Because of

Addition

- Additionally
- Furthermore
- Moreover

Attended webinar on effective academic writing in the morning.

E.g. sentence structure is very important in emphasizing the point

- A. You deserve a raise but the budget is tight
- B. The budget is tight, but you deserve a raise.

The point of this exercise is to choose the sentence that will introduce the idea that the person will be getting a raise. Placing the idea we wish to convey in the stress position helps to effectively communicate our point.

Signposting words are often used to bridge 2 different ideas to the readers. Sentence structure is very important in guiding the readers. Keep the paragraph short, number of words per sentence to about 10 to 20 and one point per sentence. It can also be used to encourage logical flow of ideas.

Active voice is encouraged in writing but it shouldn't be the only voice used. Ideally, 75% of active voice + 25% passive voice is nice.

Avoid nominalization. Strong verbs are better than the nouns. E.g. estimate > estimation. Nominalization requires accessory terms which makes the point less powerful.

We made a...

Subject

Verb

Still no idea what this
sentence is about!

...decision?

...confirmation?

...estimation?

...cake?



Ardfern / CC BY-SA 3.0

Friday - 22 April 2022

Spent some time pulling out breast cancer stage at presentation for a few countries to compare. Inspired by one of the papers that listed a few countries by income and region. Today I also learnt that there are also levels/grades to publications. Some papers are more well thought out, i.e. the paper that inspired me to pull out the stage at presentation was good, but it only listed a few examples.

Monday - 25 April 2022

Finally completed the table of breast cancer stage at presentation. Unsurprisingly, not much info for countries such as North Korea. This table has more information than other indicators though.

Tuesday - 26 April 2022

Cleaned the data for breast cancer stage at presentation. Will be looking forward to creating 8 box plots by income and region tomorrow. Hopefully, I'll be able to do it tomorrow and help Peh Joo with the chemobiome project & to also learn informed consent from Jon.

Wednesday - 27 April 2022

Cleaned up the data and managed to come up with 2 diagrams for stage at presentation. I think coding for these 2 diagrams were much faster than my previous attempts and I am able to extract more information from the plots, i.e. median, 25th percentile etc etc.

Thursday - 28 April 2022

Re-did the previous table to sort out the countries with population based and opportunistic mammography screening because there were some missing information in the previous table 1. Despite having done the table previously, some of the information was still difficult to obtain and I struggled to find the relevant information. However, this time round, I focused only on mammography in these countries and was much faster than the previous table that had CBE, BSE etc..

Friday - 29 April 2022

Met with Alya to come up with the script for the informed consent for the chemobiome project. I think working together to come up with the script helps to spot the loopholes/gaps we have in our own scripts.

Monday - 9 May 2022

Finished the table on countries with population based/opportunistic mammography screening in Asia.
Also worked a bit on the abstract.

Tuesday - 10 May 2022

Sorta done with abstract today! I learnt quite a bit about abstract submissions over zoom today. I.e. the author list, number of authors for review, what is required in an abstract (Background, Methods, Results, Conclusion) and how to structure it. My abstract focused on the relationship between income and mammography screening in Asia which is much clearer than just putting all the info I have in the abstract as there isn't a clear direction. Also learnt how to communicate with co-authors to obtain their feedback etc before the submission. I think for future reference, I will need to start working on the abstract earlier to provide the co-authors enough time to look through the draft.

Wednesday - 11 May 2022

Managed to use geom_text_repel to separate the labels in my diagram. Previously, several countries had similar data (%) for the early stage at presentation and the country code overlapped. Included _repel next to geom_text and it worked! I think the diagram looks pretty good now. Also managed to finally come up with something similar to this table for Asia! HAHAAH struggled to fill in shades for the cells initially but after fumbling around, its under format. For future reference: right click on cell -> format cell -> fill -> choose pattern & colour. Will work on the "organization" section of the table tomorrow.

Table 2 Recommendations for breast cancer screening with mammography, in order of overall healthcare spending

Country	Organization (Type)	Year	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75+
United States	US Preventive Services Task Force (A)	2016								
United States	American Cancer Society (B)	2015								*
United States ¹	American College of Obstetrics & Gynecology (C)	2017								*
United States	American College of Radiology (C)	2016								
Luxembourg	Ministry of Health (A)	NA								
Switzerland ²	League Against Cancer (B)	2016								
Norway	Cancer Registry of Norway (B)	2010								
Netherlands ³	NIPHE (A)	2017								
Germany	Federal Joint Committee (A)	2015								
Sweden ⁴	National Board of Health and Welfare (A)	2013								
Ireland	National Screening Service (A)	NA								
Austria	Austrian Cancer Aid Society (B)	2014								
Denmark	National Board of Health (A)	2014								
Belgium	Foundation Against Cancer (B)	2017								
Canada ⁵	CTFPHC (A)	2011								
Australia	Australian Government Department of Health (A)	2015								
France ⁶	National Cancer Institute (A)	2015								
Japan ⁷	National Cancer Center (A)	2016								
Iceland	Icelandic Cancer Society (B)	NA								
UK	UK National Screening Committee (A)	2012								
Finland	Cancer Society of Finland (B)	2010								
New Zealand	Ministry of Health (B)	2014								
Italy	National Screening Observatory (A)	2015								
Spain	Cancer Strategy of National Health System (A)	2009								

Recommend:	Recommend selectively:	Do not recommend	Insufficient evidence:
Every 3 years:	Every 2 years:	Every 1 year:	

Thursday - 12 May 2022

Finished filling the table with the colours. Also attempted to add in more content for the manuscript.

Friday - 13 May 2022

Submitted the abstract this morning for NCAM. Quite interesting because even during the submission I got to learn new things! I.e. co-authors, how to fill in the designation etc. Also had to re-look at the excel for countries w pop based mam screening cause the table is too wide to fit nicely into one page and is cut off. Also managed to redo my R figure hehe. Managed to add colour to the countries by region and adjust the size of the figure legends and sort of “customise” it, i.e. putting the figure legend horizontally at the top of the figure.

Tuesday - 17 May 2022

Learnt something new today. Previously, when discussing the cost effectiveness analysis for countries in Asia, I looked for countries with cost effectiveness analysis of BC and simply lumped them together as countries that deem mammography screening to be cost effective and countries that list it as ineffective. I learnt that my points were very much lacking in discussion and opinions, i.e. I was simply collecting info and listing it as it is. To “form” my own opinion, I could have mentioned “cost effectiveness of mammography screening varies in each country, etc.”, or “Mammography screening is cost effective in resource limited countries such as ...”. Discussing each country and their various reasons for mammography screening being cost effective/not could have elevated the content.

Wednesday - 18 May 2022

Tried to apply what I learnt yesterday about discussing the findings for each country to reasons why women in Asia don't screen. Instead of just listing general reasons as to why women don't screen, I tried to find reasons specific to certain countries, i.e. comparing the differences in insurance systems in Japan and South Korea, or the lack of awareness of the disease itself in China, leading to low take up rate etc. I think there is still room to talk about the low/high take up rate in certain countries, but I need to do more literature research first. I think there is space to discuss why countries in the MENA region have low screening rates etc.

Thursday - 19 May 2022

Attended a seminar for manuscript writing. Learnt about the value of having a good title, and having the variable tested, method and outcome. Chemobiome project is also gonna start next monday, have to start preparing everything. Prepared the QR code, templates for the questionnaires, prepared the logfiles with the participants' info etc. Quite excited for it to start but also lowkey nervous for the ICF ;-).

Friday - 20 May 2022

Listened to Tomiko's presentation on mother's return to workplace and it was SO cool & good!!! I think the main takeaway from her presentation is 1) presentation skills 2) how to conduct a proper review's methods and also 3) being concise in the findings. Methods: include the databases used, keywords used, and lead into the results section. Classify the findings into main points and go into each point with details. I think I learnt a lot from her presentation. Also, had to rush around to print all the forms and templates necessary for the research kits. There were more sign ups than expected so we'll have to pack more research kits next week.

Monday - 23 May 2022

Packed more kits this morning. Was quite difficult because of the details (such as the sticker on each bag), but managed to get the hang of it quite quickly. There is a lot more to packing the kits than just putting the individual items in the packet. There's a lot of printing, cutting, and pasting involved HAHAHA. Also joined Alya in her ICF and got really nervous for my ICF at night. Rushed home and did the ICF with my participant at night and I think it is not that scary, but still quite nerve-wracking. Printed a lot of the address etc and cut them out.

Tuesday - 24 May 2022

Met my first participant online yesterday night and it was honestly a bit scary AHHAHA. She was very nice and it really helped. I will be preparing to mail out the packages after tomorrow's ICF with the other participant.

Wednesday - 25 May 2022

Met the second participant and she was very involved during the ICF. She said that she wants to take a look at the ICF during lunch before sending over the signed copy.

Thursday - 26 May 2022

Prompted yesterday's participant for her ICF but there was no reply. I think she is not interested in participating. Will kiv her. Also, mailed out the first package today! Packed a lot more packages today as well.

Friday - 27 May 2022

Was supposed to meet a participant today at 1 p.m. Logged online at 12.50p.m. Participant did not show up at 1, so I waited till ~1.15. Decided to drop her an email to remind her but there was also no response. Waited online till 2 but she still did not show up. Not sure if she was a no show because she forgot/decided to withdraw. After this, I think it will be nice to send reminders the day before to prevent a no show.

Monday - 30 May 2022

Tuesday - 31 May 2022

Met participant 2 in real life today. I think it was quite funny HAHHAHA we couldn't find each other in the lobby at first because we were both younger than expected. Also, she raised a question during ICF: "Does menstruation affect the validity of the sample", which helped us prepare better for future ICF.

Wednesday - 1 June 2022

Finished going through the review once Left with diagrams and tables and also conclusion. Hope to be able to send it out to journals soon. Will also need to start working on endnote & the references ;-;

Thursday - 2 June 2022

Attended the webinar and learnt about submission to journals. Just in time HAHAAH. I think it was very informative and it helped a lot since I have to look into journal submission soon. Also, my endnote does not work ;-; got to start looking into other free reference softwares

Friday - 3 June 2022

Tried out the references part and conclusion but there's still more to do ;-;

Monday - 6 June 2022

Packed more kit for the whole week since I am meeting 2 participants over the weekend and need to prepare more kits for them. Also, went for the table tennis match in the afternoon HAHAAH was honestly very fun!

Tuesday - 7 June 2022

Attended the microbiome symposium today. Honestly thought I will not be able to understand but the speakers made it very easy to digest. Also, I quite enjoyed myself even though it was a very very long symposium HAHHAHA. I think the speakers also opened my eyes to a lot of other microbiome related projects, not just the gut. E.g. microbiome and urbanization/upcycling of food etc.

Wednesday - 8 June 2022

Thursday - 9 June 2022

Went for the ecosySTEM talk today and I went in not knowing what to expect. I thought it was gonna be similar to an Adam Khoo workshop and was kind of nervous, but the talks by the guest speakers were really really helpful. I enjoyed the sharing by the panellists very very much. I think a very common problem shared by all the wonderful females there was impostor syndrome. As much as it is a common problem, I don't think the problem resolves on its own and that it requires some courage to face this problem head on. The sharings by the guest speakers really helped a lot in making sense of the situation (feeling like an imposter) and how to tackle the problem. I think there was also a sharing on unhelpful thinking biases and what they look like. Being aware of the biases made it easier to break the pattern. Also, managed to get new connections over lunch! The ladies are super nice and willing to share their stories. It was very wholesome heh.

The 9 unhelpful thinking biases are:

- 1) Labelling
- 2) Personalisation
- 3) Mind-reading
- 4) Generalising
- 5) Black & White thinking
- 6) Shoulds and Musts
- 7) Exaggeration
- 8) Minimising
- 9) Negative filter

Friday - 10 June 2022

There was a sharing today on how to write grant proposals. I learnt that the biggest difference between a grant proposal and a cover letter is that they have different readers and different purposes. Thus, the writing style and content will be different. Writing to an editor would require more emphasis on the article and the context, but writing to a grantor will require more emphasis on how the study is significant to the industry/the company. Also, learnt how to write a good 'summary' for a person. Instead of merely describing their achievements etc, it is more impactful to bring in some personal touches. E.g. the person's name does not have to be mentioned in the intro, instead, a description of the person can be used.

Monday - 13 June 2022

Went to the lab for the first time today!!! Honestly, no lab work stuff, just labelling the samples, but it was very nice to step into a lab again heh. Received 5 packages and labelled them accordingly. Also started looking for journals that allow >6000 words for review.

Tuesday - 14 June 2022

Yesterday, I learnt that I can get “inspirations” for the journals that allow >6000 words by consolidating the journals from which I obtained my references. Spent half the day consolidating all my references into one document and also created an excel sheet to see the journals they were from. Generally, there are a few main journals that my sources were from, but only a few accept >6000 words

Wednesday - 15 June 2022

Sorted through the journals to find the ones that allow for >6000 words. I think only 3 journals, MDPI Diagnostics, Elsevier, and JNCI allow for >6000 words. Ranking them by impact factor, I think I will try for JNCI first, then Elsevier, than MDPI Diagnostics

Thursday - 16 June 2022

Checked in for the review. Gonna start working on the references and the figures. I think it is more or less done, just requires a lot of cleaning up and adding in the references. Also, realised that the references for the table are not included properly. Added the table references into the text. Started references but I am unable to access EndNote HAHHAHA. Will try out Mendeley tomorrow.

Friday - 17 June 2022

Started on references. Tried using Mendeley but I am unable to use it for Microsoft Word. Looked into using EndNote but it is not working very well. Loaded the references I created on Mendeley into Endnote after a few attempts. It is working out quite well. Have quite a few references to add in manually. Also, went for the GIS party!

Monday - 20 June 2022

Added in the references for the figures. Mendeley + EndNote still kinda wonky. Cannot add references directly in EndNote because the search keeps failing. Had to add it in Mendeley and load it into the EndNote library.

Tuesday - 21 June 2022

Reformatted all of the references and checked. Yesterday, when adding the references in, I felt that some references looked weird (the numbers were not correct). Re-did the entire references and checked to make sure it is in order. Not done w the references, have to load all of them together into EndNote.

Wednesday - 22 June 2022

Finally done with all the references, including the more difficult ones. Today I learnt how to cite NCTs. Apparently, there is a feature on EndNote that simplifies this process and that I have to download the 'filter' for this to use.

Thursday - 23 June 2022

Tried writing a cover letter. I found a template online. I tried including the methods but I think it looks kinda weird. I tried to include the significance of the study, but it is very difficult to set a 'limit' on what to include and what not to include. Tried to apply what I learnt during the 2 day talk. I think it has significance, but it is not very impactful. Today I also learnt that as much as my abstract should contain info from the paper, it shouldn't be sentences taken from different paragraphs to make up the whole abstract. Rather, it should be a sentence to summarise what the whole paragraph is talking about.

Friday - 24 June 2022

Tried to rewrite the abstract & conclusion. Also, there was a party going on and I got to meet the whole SGBCC team! Quite fun to interact with them and I love that they are a very youthful bunch of people despite being fairly experienced. I think it was great to have been able to meet them once before my internship ended.